

ANDRÉA FARIA ANDRADE

**INTEGRAÇÃO DE VARIÁVEIS ESPECTRAIS E FORMA NA
CLASSIFICAÇÃO DE IMAGENS DE ALTA RESOLUÇÃO
UTILIZANDO REDES NEURAIS ARTIFICIAIS**

Dissertação apresentada como requisito parcial à obtenção do grau de Mestre em Ciências Geodésicas, Curso de Pós-Graduação em Ciências Geodésicas, Setor de Ciências da Terra, Universidade Federal do Paraná.

Orientador: Jorge A. S. Centeno

CURITIBA

2003


**“INTEGRAÇÃO DE VARIÁVEIS ESPECTRAIS E FORMA NA
CLASSIFICAÇÃO DE IMAGENS DE ALTA RESOLUÇÃO
UTILIZANDO REDES NEURAS ARTIFICIAIS”**

POR

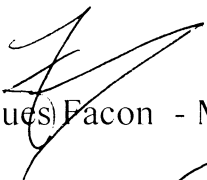
ANDREA FARIA ANDRADE

Dissertação nº 166 aprovada como requisito parcial do grau de Mestre
no Curso de Pós-Graduação em Ciências Geodésicas da Universidade
Federal do Paraná, pela Comissão formada pelos professores:

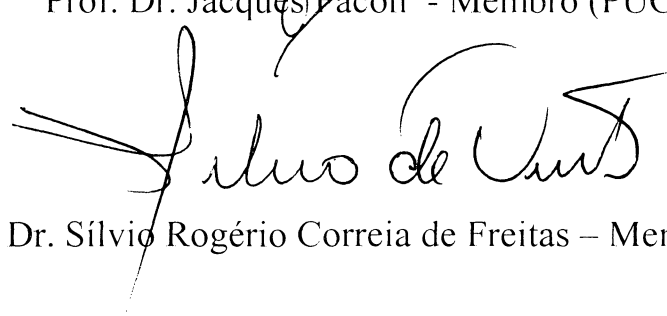
Curitiba, 26 de março de 2003.



Prof. Dr. Jorge Antonio Silva Centeno (UFPR)
Orientador e Presidente



Prof. Dr. Jacques Facon - Membro (PUC-PR)



Prof. Dr. Sílvia Rogério Correia de Freitas – Membro (UFPR)

*“Fazei tudo por Amor. Assim não há
coisas pequenas: tudo é grande.
A perseverança nas pequenas
coisas, por amor, é heroísmo”.*
(Josemaria Escrivá).

AGRADECIMENTOS

À Deus.

Ao meu marido Sandro, pela ajuda e companheirismo nas horas mais difíceis e pelo apoio e incentivo em todos os momentos.

À minha família, meus pais Edward e Cleide e minhas irmãs Adriana e Fernanda, pelo amor, apoio e incentivo.

Ao meu Orientador, Prof. Jorge Centeno, pelo apoio, dedicação e ensinamentos. E pela disposição em ajudar sempre.

Ao colega do Curso Mosar Faria Botelho, pela ajuda junto ao *MATLAB*..

Aos amigos e colegas do Curso de Pós Graduação, Andréa Tedesco, Selma, Madalena, Tatyana, Eliana, Claudinha, Moisés, Reginaldo e Jaime, pela ajuda mesmo que psicológica durante toda esta dissertação.

Aos amigos da Embap, em especial à Ângela, Claudine e Lina, pela força e apoio prestados em todos os momentos.

Aos amigos Rodrigo Barneche e Silvana Camboim, pelas ajudas prestadas.

À CAPES, pelo auxílio financeiro.

E a todos que de alguma forma contribuíram para a realização deste trabalho.

SUMÁRIO

LISTA DE FIGURAS	v
LISTA DE TABELAS	viii
RESUMO	01
ABSTRACT	02
CAPÍTULO 1 - INTRODUÇÃO	03
1.1 CONSIDERAÇÕES GERAIS	03
1.2 OBJETIVOS.....	06
1.3 JUSTIFICATIVA	07
1.4 ESTRUTURA DO TRABALHO	10
CAPÍTULO 2 – REVISÃO BIBLIOGRÁFICA	11
2.1 SENSORIAMENTO REMOTO.....	11
2.1.1 PROCESSO FUNDAMENTAL DO SENSORIAMENTO REMOTO	11
2.1.2 SISTEMAS Sensores	12
2.1.3 Imagem Digital.....	13
2.1.3.1 Resolução das Imagens Digitais	15
2.1.4 SISTEMA IKONOS	17
2.2 PROCESSAMENTO DIGITAL DE IMAGENS	18
2.2.1 PROPRIEDADES DE UMA Imagem Digital	19
2.3 RECONHECIMENTO DE PADRÕES	21
2.3.1 CLASSIFICAÇÃO DE IMAGENS.....	22
2.3.1.1 Fase de Treinamento do Classificador	25
2.3.1.2 Fase de Destino	26
2.3.1.3 Avaliação da Classificação	27
2.3.2 SEGMENTAÇÃO DE IMAGENS.....	28
2.3.2.1 Técnicas de Segmentação	29
2.3.2.1.1 Detecção de Bordas	29
2.3.2.1.2 Limiarização (Thresholding)	30

2.3.2.1.3 Segmentação Orientada a Regiões	31
2.3.2.1.3.1 Crescimento de Regiões por Agregação de Pixels	32
2.3.3 REPRESENTAÇÃO E DESCRIÇÃO DE ELEMENTOS DA IMAGEM.....	34
2.3.3.1 Textura	34
2.3.3.2 Forma	35
2.3.3.2.1 Área.....	37
2.3.3.2.2 Perímetro	37
2.3.3.2.3 Coeficiente de Compacidade	38
2.3.3.2.4 Assinaturas.....	38
2.3.3.2.5 Relação Eixo Menor e Maior da Elipse.....	41
2.3.3.2.6 Código da Cadeia	42
2.3.3.2.7 O fecho convexo	43
2.4 REDES NEURAIIS	44
2.4.1 O SISTEMA NERVOSO.....	45
2.4.1.1 O Neurônio Biológico.....	46
2.4.1.2 Redes Neurais Artificiais – O Modelo MCP (McCulloch e Pitts).....	47
2.4.1.2.1 Funções de Ativação	50
2.4.1.3 Arquitetura das Redes Neurais.....	51
2.4.1.4 Conjunto de Dados.....	53
2.4.1.5 Aprendizado de uma RNA.....	54
2.4.1.5.1 Aprendizado Supervisionado	54
2.4.1.5.2 Aprendizado Não-Supervisionado	55
2.4.1.6 Treinamento	56
2.4.1.6.1 O Algoritmo Backpropagation.....	57
2.4.1.7 Exemplos de Aplicação das Redes Neurais	57
CAPÍTULO 3 – MATERIAIS E MÉTODOS	59
3.1 MATERIAIS EMPREGADOS.....	59
3.1.1 IMAGEM UTILIZADA	59
3.2 METODOLOGIA	60
3.2.1 SEGMENTAÇÃO DA IMAGEM.....	60

3.2.2 CODIFICAÇÃO DE CADA SEGMENTO DA IMAGEM	62
3.2.3 DEFINIÇÃO DOS PARÂMETROS NECESSÁRIOS À CLASSIFICAÇÃO POR REDES NEURAS ARTIFICIAIS	64
3.2.3.1 Parâmetros da Arquitetura da Rede	64
3.2.4 CLASSIFICAÇÃO ORIENTADA AO OBJETO UTILIZANDO REDES NEURAS ARTIFICIAIS	68
3.2.5 AVALIAÇÃO DA CLASSIFICAÇÃO	76
3.2.6 COMPARAÇÃO DO RESULTADO EM RELAÇÃO À CLASSIFICAÇÃO <i>PIXEL A PIXEL</i> POR REDES NEURAS E MÉTODO DA MÁXIMA VEROSSIMILHANÇA	77
3.2.7 COMPARAÇÃO DA CLASSIFICAÇÃO <i>ORIENTADA AO OBJETO</i> POR REDES NEURAS ARTIFICIAIS E PELO MÉTODO DA DISTÂNCIA MÍNIMA DO <i>SOFTWARE ECOGNITION</i>	78
3.2.8 CLASSIFICAÇÃO DE OUTRA ÁREA DA IMAGEM <i>IKONOS</i> DE MATINHOS	79
CAPÍTULO 4 – RESULTADOS E DISCUSSÕES.....	82
4.1 SEGMENTAÇÃO DA IMAGEM	82
4.2 CLASSIFICAÇÃO <i>PIXEL A PIXEL</i> POR REDES NEURAS E MÉTODO DA MÁXIMA VEROSSIMILHANÇA.....	84
4.3 CLASSIFICAÇÃO <i>ORIENTADA AO OBJETO</i> UTILIZANDO REDES NEURAS ARTIFICIAIS	87
4.4 CLASSIFICAÇÃO <i>ORIENTADA AO OBJETO</i> POR REDES NEURAS ARTIFICIAIS E ATRAVÉS DO <i>SOFTWARE ECOGNITION</i>	94
4.5 VERIFICAÇÃO DA EFICIÊNCIA DA REDE NEURAL EM OUTRA ÁREA	96
4.6 ESCOLHA DO MELHOR PARÂMETRO DE FORMA.....	101
CAPÍTULO 5 – CONCLUSÕES E RECOMENDAÇÕES.....	103
REFERÊNCIAS BIBLIOGRÁFICAS.....	107

LISTA DE FIGURAS

FIGURA 01 - Componentes de um Sistema de Sensoriamento Remoto	11
FIGURA 02 - Sistema de Aquisição de Informação	14
FIGURA 03 - (a) imagem com 2 níveis de cinza; (b) imagem com 32 níveis de cinza	16
FIGURA 04 - Conceitos de vizinhança-de-4, vizinhança diagonal e vizinhança-de-8	19
FIGURA 05 - Limiarização de uma Imagem Monocromática Utilizando Limiar T : (a) Histograma Original, (b) Histograma da Imagem Binarizada.....	31
FIGURA 06 - (a) Imagem Original Usando um Ponto Semente; (b) Estágio Primário de Crescimento de uma Região; (c) Estágio Intermediário de Crescimento de uma Região; (d) Estágio Final	33
FIGURA 07 – Fronteira Simples e sua Correspondente Assinatura	39
FIGURA 08 - Direções do (a) Código da cadeia direcional-de-4 e (b) Direcional-de-8	42
FIGURA 09 - Representação da Forma Através do Código da Cadeia usando 4 Direções	42
FIGURA 10 - Representação do Código da Cadeia Normalizado	43
FIGURA 11 - Região (s) e sua Deficiência Convexa Representada como Áreas na Cor Cinza	44
FIGURA 12 - Componentes de um Neurônio Biológico	47
FIGURA 13 - Neurônio de McCulloch e Pitts	48
FIGURA 14 - Algumas Funções de Ativação	51
FIGURA 15 - Exemplo de Uma Rede Neural	52
FIGURA 16 - Aprendizado Supervisionado.....	55
FIGURA 17 - Aprendizado Não-Supervisionado.....	56
FIGURA 18 - Imagem <i>IKONOS</i> de Matinhos.....	60

FIGURA 19 – Imagem Resultante da Segmentação no <i>Software eCognition</i>	62
FIGURA 20 - Imagens de Segmentos Codificados	63
FIGURA 21 - Imagem Sintética	67
FIGURA 22 - Imagem <i>IKONOS</i> Simples.....	67
FIGURA 23 - Esquema da Montagem dos Programas de Coleta de Amostras	72
FIGURA 24 - Segunda Imagem de Matinhos Utilizada na Classificação por Redes Neurais Artificiais	80
FIGURA 25 - Imagem Segmentada em Parâmetros de Escala Diferentes. (A) Imagem Segmentada com Parâmetro de Escala De 25. (B) Imagem Segmentada com Parâmetro de Escala de 60. (C) Imagem Segmentada com Parâmetro de Escala de 50	83
FIGURA 26 - Resultado da Classificação Utilizando o Método da Máxima Verossimilhança (<i>ENVI</i>)	84
FIGURA 27 - Resultado da Classificação Utilizando o Método das Redes Neurais Artificiais (<i>ENVI</i>).....	85
FIGURA 28 - Esquema da Arquitetura da Rede Utilizada.....	88
FIGURA 29 - Resultado das Classificações Puramente Espectral e com os Índices de Forma <i>compactidade + razão eixos + 1º momento</i>	93
FIGURA 30 - Classificação Realizada no <i>Software ECOGNITION</i>	95
FIGURA 31 - Classificação Realizada pelo Método das Redes Neurais Artificiais.	95
FIGURA 32 - Imagem Original.....	96
FIGURA 33 - Imagem Original.....	97
FIGURA 34 - Imagem Resultante da Classificação Utilizando Parâmetros Espectrais e Forma, Através do <i>Software ECOGNITION</i>	98
FIGURA 35 - Imagem Resultante da Classificação Utilizando os Mesmos Parâmetros de Arquitetura da Rede da Imagem Utilizada Anteriormente	98
FIGURA 36 - Imagem Resultante da Classificação Puramente Espectral (<i>IVP-R-G</i>), Utilizando as Redes Neurais Artificiais.....	99
FIGURA 37 - Imagem Resultante da Classificação com Novo Treinamento para a	

Rede, Utilizando a Melhor Combinação dos Parâmetros de Forma Encontrados para a Imagem Anterior	100
FIGURA 38 – Acurácia Global em Função do Número de Parâmetros de Forma ...	101

LISTA DE TABELAS

TABELA 01 - Principais Características do Sensor <i>IKONOS II</i>	17
TABELA 02 - Coeficiente <i>Kappa</i> e o Desempenho da Classificação	28
TABELA 03 - Tabela de Entrada das Amostras de Treinamento, Apresentada à Rede	71
TABELA 04 - Tabela de Amostras de Treinamento com Alguns Parâmetros de Forma	73
TABELA 05 - Variação do Número de Parâmetros de Entrada Apresentados à Rede ao Longo da Pesquisa	74
TABELA 06 - Variação do Número de Parâmetros de Entrada Apresentados à Rede ao Longo da Pesquisa	75
TABELA 07 - Variação do Número de Parâmetros de Entrada Apresentados à Rede ao Longo da Pesquisa	76
TABELA 08 - Matriz de Confusão Resultante da Classificação Através do Método das Redes Neurais Artificiais (<i>ENVI</i>).....	86
TABELA 09 - Matriz de Confusão Resultante da Classificação Através do Método da Máxima Verossimilhança (<i>ENVI</i>).....	87
TABELA 10 - Resultado das Classificações Realizadas com os Índices de Forma Inseridos <i>um a um</i>	89
TABELA 11 - Resultado das Classificações Realizadas com os Índices de Forma Inseridos <i>dois a dois</i>	89
TABELA 12 - Resultado das Classificações Realizadas com Três Índices de Forma Adicionadas aos Parâmetros Espectrais.....	90
TABELA 13 - Resultado das Classificações Realizadas com Quatro Índices de Forma Adicionadas aos Parâmetros Espectrais.....	91
TABELA 14 - Resultado das Classificações Realizadas com Cinco Índices de Forma Adicionadas aos Parâmetros Espectrais.....	91

TABELA 15 - Resultado das Classificações Realizadas com Seis a Nove Índices de Forma Adicionadas aos Parâmetros Espectrais	92
--	----

RESUMO

A classificação de imagens tem sido um importante processo para a produção de mapas temáticos, e uma das principais fontes de dados para os Sistemas de Informações Geográficas (SIG). Com o surgimento dos sensores de alta resolução espacial, as aplicações não serão mais restritas aos estudos de clima e monitoramento de florestas, mas pode ser estendido para outras áreas onde há a necessidade de uma descrição mais detalhada da superfície terrestre, como telecomunicações, planejamento urbano, mapas de arruamentos e cadastro e planejamento de traçados de rodovias, ferrovias e dutos. No entanto, para o completo aproveitamento das novas imagens, com melhor resolução espacial, abaixo de 5 metros, torna-se necessário disponibilizar de novas ferramentas, que se estendam além da análise dos *pixels* isoladamente, possibilitando a modelagem de objetos, levando em consideração outros atributos da imagem, além dos espectrais. Na classificação *orientada ao objeto* aplicada a imagens de alta resolução espacial, a imagem é primeiramente segmentada e posteriormente os segmentos, que correspondem a objetos, ou parte deles são classificados usando informações espectrais, espaciais, como a forma ou a orientação dos segmentos, e a textura. O presente trabalho tem como objetivo a inserção da forma no processo de classificação, por considerar que ela é relevante na identificação de objetos presentes na imagem. Para isto, o efeito de vários parâmetros de forma é avaliado usando uma imagem urbana *IKONOS*, com resolução espacial de 1 metro. O algoritmo de classificação se baseia no princípio das Redes Neurais Artificiais. A utilização de Redes Neurais Artificiais emerge como uma alternativa vantajosa aos classificadores baseados em conceitos estatísticos, uma vez que nenhuma hipótese prévia sobre a distribuição dos dados a serem classificados é exigida. A análise de forma comprovou ser uma poderosa ferramenta para a discriminação de objetos que possuem a mesma aparência espectral. Os resultados são comparados com métodos de classificação convencionais baseados em dados puramente espectrais, para fins de avaliação do ganho obtido no processo, incluindo a forma.

ABSTRACT

Remote sensing image classification has been an important step for the production of thematic maps, and one of the main data sources for the geographical information systems (GIS). With the development of high spatial resolution sensors, the applications of remote sensing are no more restricted to climate studies and environmental monitoring, but can be extended to other areas that need a more detailed description of the earth's surface, such as telecommunication and urban planning. However, for the complete use of the higher spatial resolution images, new tools are necessary, that enable to extend the analysis further than the pixel based classification. Approaches like the object oriented image analysis enable the modeling of objects, taking into account other attributes of the image than the spectral ones. In the object oriented classification approach, the image is segmented and later the segments that correspond to objects, or parts of them, are classified using spectral and spatial attributes, as the form or the orientation of the segments, and the texture. The present work deals with the insertion of shape information in the classification process, since it is important for identifying the objects in the image. For this purpose, the effect of the inclusion of several shape parameters into the classification schema is evaluated, using an urban *IKONOS* image, with spatial resolution of 1 meter. The classification algorithm is based on the neural network approach. The use of artificial neural nets emerges as an advantageous alternative to the statistical classifiers, since any previous hypothesis about the distribution of the data is demanded. The analysis of shape is a powerful tool for the discrimination of objects that possess the same spectral appearance. The results are compared to the results of a pixel based spectral classification, in order to evaluate the contribution of the inclusion of shape parameters in the process. A comparison to other classification algorithms is also presented.

CAPÍTULO 1

INTRODUÇÃO

1.1 CONSIDERAÇÕES GERAIS

Uma das grandes vantagens do Sensoriamento Remoto é a facilidade que as imagens obtidas oferecem para expandir a capacidade de percepção visual do ser humano, pela medição da energia eletromagnética em regiões do espectro imperceptíveis a olho nu. O registro de informações espectrais em várias bandas do espectro eletromagnético facilita a interpretação e classificação da cobertura da Terra através das imagens. O processo de classificação de imagens consiste na atribuição de cada *pixel* a uma classe pertencente a um grupo de classes, e grande parte desta tarefa é realizada através da utilização das propriedades espectrais dos objetos presentes na cena. Esta análise pode ser visual ou automática, usando computadores, e é realizada estudando-se a variação espacial ou temporal das propriedades espectrais da superfície dos objetos. A análise visual das imagens, no entanto, não se restringe às propriedades espectrais. Entre outros aspectos que podem ser explorados encontram-se a textura, o tamanho e a forma dos objetos, bem como a associação espacial entre os objetos que compõem a cena.

Quando a classificação é feita usando computadores, geralmente são utilizados classificadores baseados em princípios estatísticos, pois estes são os que habitualmente encontram-se implementados na maioria dos programas utilizados em processos de extração de informações de imagens provenientes de sensoriamento remoto. Ainda, a maioria dos programas disponíveis são classificadores *pixel a pixel*, ou seja, analisam a resposta espectral de um pixel isoladamente para tomar uma decisão a respeito da classe mais provável.

Como a resolução espacial das imagens de satélite é limitada pelo sensor

utilizado, as aplicações do Sensoriamento Remoto ficaram restritas há até alguns anos aos estudos de clima, mineração, controle de queimadas, monitoramento de florestas, etc., ou seja, aplicações que abordavam grandes áreas. Com o avanço da tecnologia dos sensores e o surgimento dos sistemas de sensoriamento remoto de alta resolução espacial, como o sistema *IKONOS II*, lançado em meados 1999, há a possibilidade da identificação de ruas e casas, fato que vem provocando uma verdadeira revolução nos processos de geração e manutenção de bases cartográficas, como também nos processos clássicos de classificação de imagens. Com isto, a eficiência dos algoritmos de classificação espectral passou a ser questionada e novos métodos que ajudem na extração de informações vêm sendo analisados.

Um dos fatores que mais ajuda na discriminação e descrição de objetos, além das características espectrais, é a forma. Esta constatação é atualmente verificada em programas de análise de imagens de sensoriamento remoto. Por exemplo, na abordagem "orientada a objeto", a imagem é segmentada e posteriormente os segmentos classificados segundo suas características espectrais e outros descritores espaciais e de textura. No entanto, existe uma grande variedade de parâmetros de forma na literatura, pois a forma, por ser um conceito subjetivo desenvolvido na linguagem humana, é difícil de ser parametrizada em uma imagem no formato "raster". Mais ainda, é necessário levar em consideração que determinados parâmetros podem ser apropriados para um tipo de objeto e falhar quando são aplicados na identificação de outros objetos em outras imagens, devido à diferente resolução espacial e a natureza dos objetos.

A integração de dados de diferentes naturezas, espectrais e espaciais, requer o uso de uma ferramenta apropriada. A utilização de Redes Neurais Artificiais emerge como uma alternativa vantajosa aos classificadores baseados em conceitos estatísticos, uma vez que nenhuma hipótese prévia sobre a distribuição dos dados a serem classificados é exigida e os dados de entrada podem ter escalas diferentes, ou seja, os dados podem apresentar em escalas de dados reais, inteiros, escalas logarítmicas, etc..

Com a capacidade intrínseca de aprender, as redes neurais surgiram como uma ferramenta poderosa na área de reconhecimento de padrões. Sua capacidade de fazer suposições mais sutis a respeito da distribuição dos dados de entrada do que os métodos estatísticos tradicionais e a capacidade de formar fronteiras de decisão não-lineares no espaço de características, têm aumentado o seu uso significativamente.

A flexibilidade dos algoritmos de classificação baseados em redes neurais permite integrar variáveis diferentes na identificação de objetos, como por exemplo de características espectrais e espaciais. A análise de forma e tamanho é uma poderosa ferramenta para a discriminação de objetos que possuem a mesma aparência espectral. Considerando, por exemplo, duas regiões de uma imagem ocupadas por construções e asfalto, na maioria das vezes ambas apresentam aparência espectral muito similar, de maneira que sua discriminação não seja possível. Isto, no entanto, pode se tornar mais fácil se for levado em consideração que construções possuem uma forma retangular, ao contrário do asfalto que já possui uma forma mais alongada.

Assim como a análise espectral baseia-se nos valores digitais das bandas, é necessário dispor de parâmetros para descrever os atributos espaciais dos objetos. Para analisar a forma é necessário conhecer os limites dos objetos nas imagens, algo que a princípio é desconhecido. Por isto, os algoritmos de análise de forma partem de uma imagem temática, onde os objetos se encontrem mais definidos. A imagem temática pode ser o resultado de uma classificação ou a segmentação da imagem.

Entre os parâmetros utilizados como descritores da variação espacial do objeto encontram-se: Área (A), Perímetro (P), Coeficiente de Compacidade (CC), número de formas e teoria dos fractais, através da análise da dimensão fractal (D). Neste trabalho alguns destes parâmetros serão analisados com o objetivo de se obter os mais relevantes ao auxílio no processo de classificação.

1.2 OBJETIVOS

O objetivo principal deste trabalho é desenvolver um método para a classificação automática de objetos presentes em imagens de alta resolução espacial, utilizando para isto além da informação espectral contida nos pixels, a informação derivada da análise da forma.

Como objetivos específicos tem-se:

- a) avaliar este método em imagens de sensoriamento remoto de alta resolução espacial, na identificação de objetos como edificações e arruamentos, usando dados *IKONOS*;
- b) avaliar diferentes descritores de forma;
- c) comparar diferentes descritores de forma;
- d) avaliar o potencial das redes neurais artificiais para a integração de informação espacial e espectral na classificação.

1.3 JUSTIFICATIVA

O planejamento da expansão urbana no Brasil, de uma forma geral, não tem conseguido acompanhar o aumento acentuado da população das cidades e todo o conjunto de necessidades que esta situação implica, tendo como conseqüências o estabelecimento de um quadro de crescimento desordenado em seu espaço territorial. As conseqüências deste processo inadequado de crescimento são comuns em todas as cidades grandes: falta de condições sanitárias mínimas em muitas áreas; ocupação de áreas inadequadas; destruição de recursos de valor ecológico; poluição do meio ambiente; habitações em condições precárias de vida. Tudo isso se reflete na qualidade de vida urbana, repercutindo diretamente sobre a saúde das pessoas (LIMA e PHILIPS, 2000).

Esta deficiência, entre outros motivos, é decorrente também da falta de conhecimento rigoroso do espaço físico de seu território, o qual, tecnicamente pode ser superado através da organização de um cadastro técnico polivalente elaborado dentro de uma concepção multidisciplinar que o assunto requer.

De acordo com ELSTNER (1971), o cadastro deve servir para múltiplos fins, como:

- base para a cobrança racional de impostos;
- garantir a posse da propriedade;
- base para um melhoramento do registro de terras;
- reforma agrária;
- base para projetos de desenvolvimento: construção de estradas, projetos de economia hidrográfica (irrigação, drenagem, plantas hidroelétricas, águas subterrâneas, etc.), planejamento de novos povoados, planejamento para o desenvolvimento urbano, base para a geografia regional, base para planejamento na remodelação das propriedades (para construções, ruas, etc.);
- base para a manutenção atualizada dos mapas topográficos básicos e gerais.

O conhecimento do espaço físico pode ser atingido obtendo informações com Sensoriamento Remoto, pois existe uma grande variedade de sensores disponíveis na atualidade. Com o surgimento dos sistemas de sensoriamento remoto de alta resolução espacial, as aplicações não são mais restritas aos estudos de clima, controle de queimadas e monitoramento de florestas, recursos hídricos, mas sim na elaboração de mapas urbanos, mapas de arruamentos e cadastro, elaboração de mapas de uso e ocupação do solo, no anteprojeto de estradas, na localização de caminhos e auto-estradas, na agricultura, com o monitoramento da variabilidade espacial das culturas, e no Sistema de Informações Geográficas (SIG) através de redes, telecomunicações, planejamento, meio ambiente.

Atualmente verifica-se uma verdadeira revolução na área de sensores abordo de satélites de sensoriamento remoto. Além dos satélites das séries Landsat e Spot, um amplo conjunto de novas tecnologias de imageamento já está em órbita ou em planejamento. Os sensores hiperespectrais, com grande número de bandas de largura de faixa estreita, como o *MODIS*, que conta com 39 bandas, lançado em 1998, ou os de alta resolução espacial, como o *IKONOS II* (1 m) e o *QUICKBIRD* (61 cm), têm revolucionado os paradigmas da Fotogrametria e têm tornado viável o uso de imagens orbitais para a aplicação cadastral, principalmente na extração automática de feições (linhas e áreas) úteis na atualização de cartas e na segmentação e classificação de áreas urbanas.

O recente aumento da resolução espacial dos sensores comercialmente disponíveis tornou interessante o uso de imagens orbitais para fins de mapeamento do espaço urbano. No entanto, são necessários aprimoramentos nas ferramentas de análise e processamento de imagens para adaptá-las ao conteúdo de informação das novas imagens.

O processamento de imagens é certamente uma área em crescimento. Diversos temas científicos são abordados e em alguns casos de caráter interdisciplinar. Entre eles pode-se citar a compressão de imagens, a análise em multiresolução, análise

estatística entre outras. A classificação de imagens de sensoriamento remoto tem sido um importante processo para a produção de mapas temáticos e uma das principais fontes de dados para os sistemas de informações geográficas, sendo que a mesma tem sido aperfeiçoada em decorrência do avanço da tecnologia dos novos sensores.

De acordo com MEYER e WERTH (1990, p.10-13), cerca de 95% de toda classificação procura realizar a tarefa de fotointerpretação (feita por um analista com inúmeras variáveis), usando apenas uma variável, os atributos espectrais. Portanto não é surpreendente que hajam erros em mapas temáticos provenientes da classificação. Vários cientistas sabendo desta condição tentaram aumentar a acurácia e a qualidade da classificação derivada do sensoriamento remoto através da incorporação de dados auxiliares no processo de classificação (por exemplo, MASON *et al.*, 1988, p. 195-212; HUTCHINSON, 1982, p.123-130; FRANKLIN e WILSON, 1992, p.449-454).

Geralmente, na análise de imagens de satélite, o processamento é feito pixel a pixel, utilizando apenas seu atributo espectral. Outra opção é dividir a imagem em regiões e extrair os atributos espaciais, geométricos, ou seja, a informação espacial que envolve a relação entre os pixels e seus vizinhos, e utilizá-los para posterior classificação destas regiões.

A análise de forma é uma poderosa ferramenta para a discriminação de objetos que possuem a mesma aparência espectral. Assim como a análise espectral baseia-se nos valores digitais das bandas, é necessário dispor de um parâmetro de forma para descrever objetos. Para analisar a forma se faz necessário conhecer os limites dos objetos nas imagens, algo que a princípio é desconhecido. Para isto, os algoritmos de análise de forma partem de uma imagem temática, se possível binária, onde os objetos se encontrem mais definidos. A imagem temática pode ser o resultado de uma classificação supervisionada ou da segmentação da imagem.

Tarefas como classificação de padrões são executadas pelo cérebro humano de forma rotineira, porém os sistemas computacionais tradicionais ainda estão longe de alcançar um desempenho similar. Na atualidade, busca-se emular o raciocínio humano

através de algoritmos de computador com a finalidade de extrair informações das imagens automaticamente, tentando simular a análise visual humana. Para isto, técnicas de Inteligência Artificial vêm sendo implementadas no processo de análise de imagens. Neste contexto, as Redes Neurais Artificiais são uma das linhas da inteligência artificial que mais interesse levantam no processamento de imagens. NELSON e ILLINGWORTH (1991), definiram as Redes Neurais Artificiais (RNA's) como sendo técnicas inovadoras de processamento de informação com o objetivo básico de simular em computadores, o funcionamento do cérebro humano.

Este trabalho se justifica pela utilização das RNA's na classificação dos padrões presentes em imagens de alta resolução espacial, utilizando a integração de variáveis espectrais e espaciais (a forma dos objetos), como também a comparação dos vários parâmetros de forma, a fim de se obter uma qualidade e acurácia maior na geração de mapas temáticos oriundos da classificação.

1.4 ESTRUTURA DO TRABALHO

Esse trabalho está estruturado em cinco capítulos.

O primeiro capítulo apresenta a origem, importância, os objetivos e a estrutura do trabalho.

O capítulo 2 descreve conceitos de Sensoriamento Remoto, Processamento Digital de Imagens, Reconhecimento de Padrões e de Redes Neurais Artificiais.

No capítulo 3 é apresentada a parte metodológica do trabalho, onde são mostrados os materiais e métodos utilizados.

No capítulo 4 são apresentados os resultados obtidos e são realizadas discussões sobre os mesmos.

O capítulo 5 apresenta as conclusões e recomendações para trabalhos futuros.

Finalmente, são apresentadas as referências bibliográficas utilizadas no desenvolvimento do presente trabalho.

CAPÍTULO 2

REVISÃO BIBLIOGRÁFICA

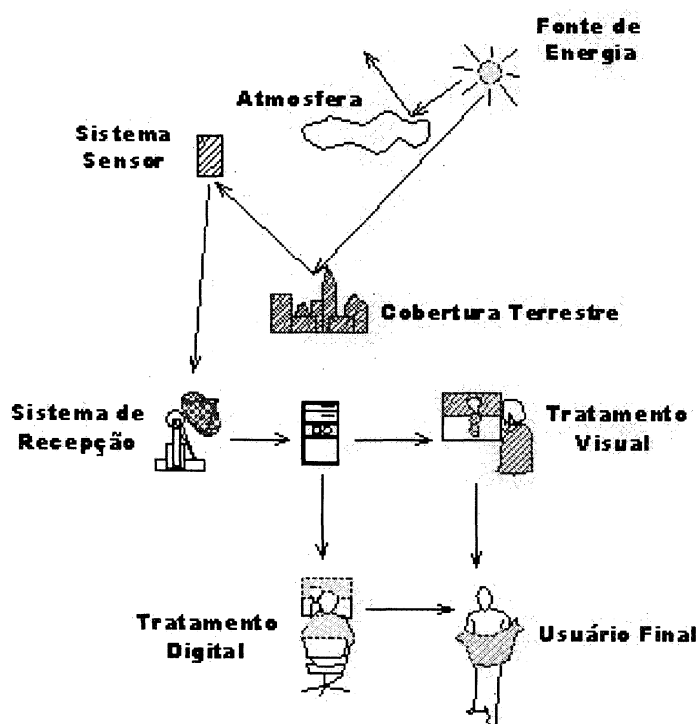
2.1 SENSORIAMENTO REMOTO

O sensoriamento remoto entende-se como o meio pelo qual dados sobre determinado objeto da superfície terrestre são obtidos através de sensores transportados por satélites ou aeronaves, sendo esta aquisição realizada sem nenhum contato físico direto com o alvo de investigação (CROSTA, 1992).

2.1.1 PROCESSO FUNDAMENTAL DO SENSORIAMENTO REMOTO

O processo fundamental do sensoriamento remoto pode ser visualizado na Figura 1.

FIGURA 1 - COMPONENTES DE UM SISTEMA DE SENSORIAMENTO REMOTO.



FONTE: ADAPTADO DE CHUVIECO, 1990, P.26.

Da análise da Figura 1 destacam-se alguns componentes da fase de aquisição de dados de sensoriamento remoto:

1. a fonte de energia eletromagnética;
2. a cobertura terrestre, ou seja, o alvo que recebe, absorve ou reflete esta radiação de acordo com suas características físicas;
3. o sistema sensor, composto pelo sensor e sua plataforma que o transporta;
4. a atmosfera por onde se propaga a energia;
5. sistema de recepção, de onde é recebida a informação transmitida pela plataforma e gravadas em um formato apropriado;
6. intérprete que analisa as informações em forma de imagens, e;
7. usuário final, que se encarrega de obter informações através da interpretação da imagem.

2.1.2 SISTEMAS SENSORES

O sensor é um sistema de aquisição de dados a respeito de uma cena (objeto físico que emite radiação). Um sensor é formado basicamente por um sistema óptico que é responsável pela formação da imagem no plano focal, onde está localizado um conjunto de detectores. Estes detectores recebem a energia radiante focalizada pela lente e a transformam em sinal elétrico.

Os detectores medem a intensidade de energia incidente. Como as propriedades espectrais dos objetos não são constantes ao longo do espectro eletromagnético, o armazenamento das interações do objeto com a radiação solar refletida pelos objetos em diferentes porções, ou bandas, do espectro eletromagnético produz diferentes imagens do objeto em questão. Como objetos diferentes possuem características espectrais diferentes, a análise das variações espectrais nas imagens permite a discriminação de objetos.

2.1.3 IMAGEM DIGITAL

O termo imagem monocromática, ou simplesmente imagem, refere-se à função bidimensional de intensidade da luz $f(x, y)$, onde x e y denotam as coordenadas espaciais e o valor de f em qualquer ponto (x, y) é proporcional ao brilho (ou níveis de cinza) da imagem naquele ponto. Uma imagem digital é uma imagem $f(x, y)$ discretizada tanto em coordenadas espaciais quanto em brilho (MASCARENHAS e VELASCO, 1989, p.1.1-1.3).

A função $f(x, y)$ representa o produto da interação entre a iluminância $i(x, y)$, que é a quantidade de luz que incide sobre o objeto, e as propriedades de reflectância próprias do objeto, que podem ser representadas pela função $r(x, y)$, cujo valor exprime a fração de luz incidente que o objeto refletir no ponto (x, y) .

$$f(x, y) = i(x, y) \cdot r(x, y) \quad (1)$$

com:

$$0 < i(x, y) < \infty \quad e \quad (2)$$

$$0 < r(x, y) < 1 \quad (3)$$

A equação (3) indica que a reflectância é limitada entre 0 (absorção total) e 1 (reflectância total).

A intensidade armazenada em uma imagem monocromática nas coordenadas (x, y) , é denominada valor digital ou nível de cinza (L) da imagem naquele ponto. Este valor está no intervalo:

$$L_{\min} \leq L \leq L_{\max} \quad (4)$$

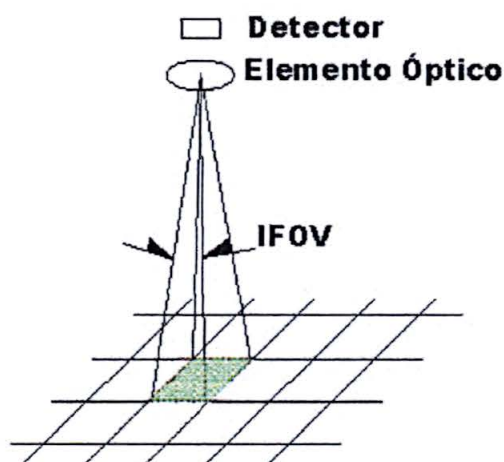
sendo L_{\min} e L_{\max} valores positivos e finitos.

O intervalo $[L_{\min}, L_{\max}]$ é denominado escala de cinza da imagem. É comum utilizar um intervalo de números inteiros $[0, W-1]$ onde $L=0$ significa pixel preto e $L=W-1$ representa pixel branco. Normalmente, W é uma potência inteira positiva de 2. Todos os valores intermediários são tons de cinza variando entre o branco e o preto (GONZALES e WOODS, 2000, p. 21).

No caso de uma imagem que possui informações em bandas distintas, é necessária uma função $f(x, y)$ para cada banda (MARQUES e VIEIRA, 1999, p. 28; GONZALES e WOODS, 2000, p.21).

De acordo com CROSTA (1992) as imagens de sensoriamento remoto são constituídas por um arranjo de elementos sob a forma de uma malha ou grade (*matriz*). Cada elemento desta matriz (*pixel*), derivado do inglês "*picture element*", tem sua localização definida de acordo com um sistema de coordenadas do tipo "*coluna e linha*", representados por "*x*" e "*y*", respectivamente. Cada pixel representa uma área que é definida pela projeção geométrica de um detector na superfície terrestre, e este possui também um atributo numérico em "*z*", que indica o nível de cinza representando a média da radiância, para a área da superfície terrestre correspondente. O ângulo de visão atribuído a um detector é denominado IFOV (*Instantaneous Field of View*), mostrado na Figura 2. Sua projeção no solo define a resolução espacial do sistema. Quanto melhor a resolução espacial, maior o nível de detalhes perceptível na imagem (FONSECA *et al.*, 2000, p.2.4).

FIGURA 2 - SISTEMA DE AQUISIÇÃO DE INFORMAÇÃO.



FONTE: ADAPTADO DE FONSECA ET AL., 2000, P.2.4.

2.1.3.1 Resolução das Imagens Digitais

Conforme CROSTA (1992), o termo resolução pode se desdobrar em quatro diferentes (e independentes) parâmetros:

- resolução espacial;
- resolução espectral;
- resolução temporal e,
- resolução radiométrica.

De acordo com CHUVIECO (1990, p. 91), a resolução espacial é o tamanho da projeção do *IFOV* na superfície observada, e a mesma possui uma relação com o nível de detalhes perceptíveis na imagem.

A resolução espectral indica o número e a largura das bandas espectrais que o sensor pode obter, ou seja das várias imagens coletadas pelo sensor com respostas espectrais diferentes. Quanto maior a resolução espectral do sensor, ou seja, quanto maior o número de canais espectrais do sensor e menor a largura destas faixas, mais fácil é a distinção dos diferentes alvos detectados.

A resolução temporal é definida pelo intervalo de tempo medido entre duas aquisições consecutivas de imagens de uma mesma região. Ou seja, refere-se à periodicidade com que o sensor adquire imagens da mesma porção da superfície terrestre (CHUVIECO, 1990, p.95).

A resolução radiométrica de um sistema é definida pelo número de bits que representam um valor digital (nível de cinza). O número de valores de níveis de cinza possíveis é dado por:

$$K = 2^m, \quad (5)$$

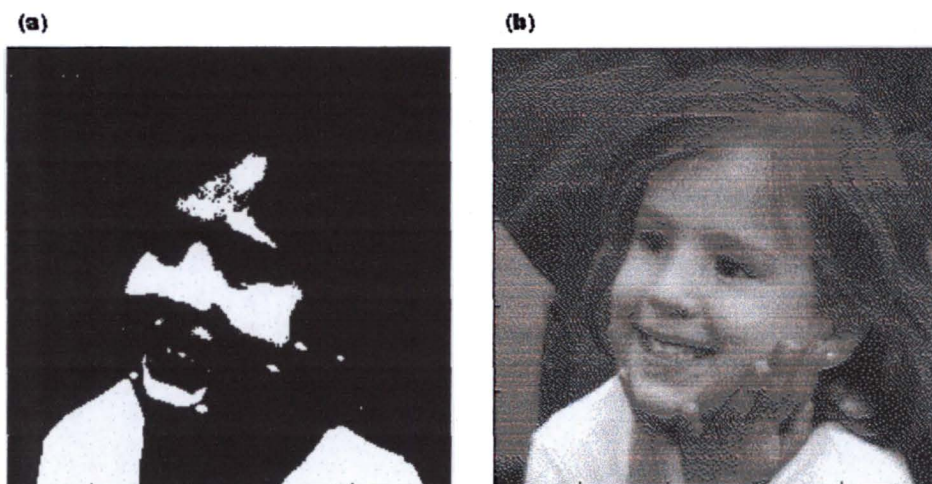
onde m é o número de bits.

Pode-se dizer que o processo de digitalização de uma imagem “contínua” ou analógica corresponde a uma discretização (processo de amostragem) da cena em observação, através da superposição de uma malha hipotética, e uma atribuição de

valores inteiros (níveis de cinza) a cada ponto dessa malha, denominado de processo de quantização (MARQUES e VIEIRA, 1999, p.29).

Comumente aos outros tipos de resolução, quanto maior for a resolução radiométrica, tanto melhor poderá ser interpretada a imagem. Para entender melhor esse conceito, pode-se pensar em uma imagem com apenas 2 níveis (preto e branco, Figura 3a) em comparação com uma imagem com 32 níveis de cinza entre o branco e o preto (Figura 3b); obviamente a quantidade de detalhes perceptíveis na segunda será maior do que na primeira e portanto a segunda imagem terá uma melhor resolução radiométrica.

FIGURA 3 – (A) IMAGEM COM 2 NÍVEIS DE CINZA; (B) IMAGEM COM 32 NÍVEIS DE CINZA.



FONTE: [HTTP://WWW.DCA.FEE.UNICAMP.BR/~AJCRUZ/LISTA01/PLISTA01.HTML](http://www.dca.fee.unicamp.br/~ajcruz/Lista01/PLISTA01.html) EM 17/03/03.

Um sensor pode ser considerado de alta resolução quando possuir uma alta resolução espacial, e/ou espectral e/ou radiométrica, como é o caso dos sensores de alta resolução espacial *IKONOS II* e *QUICKBIRD*, com resoluções de 1m e 0,61cm, respectivamente, para as bandas pancromáticas, aumentando o nível de detalhamento captado, antes apenas observável através de fotos aéreas. Um exemplo de sensor de alta resolução espectral é o sensor *MODIS*, que possui 39 bandas espectrais, permitindo uma maior penetração na atmosfera e maior poder de discriminação dos

alvos terrestres.

2.1.4 SISTEMA *IKONOS II*

O satélite *IKONOS II* foi lançado no dia 24 de Setembro de 1999 e está operando desde o início de 2000. Ele é operado pela *SPACE IMAGING* que detém os direitos mundiais de comercialização em nível mundial. As principais características deste sensor são apresentadas através da tabela 1 abaixo.

TABELA 1 - PRINCIPAIS CARACTERÍSTICAS DO SENSOR *IKONOS II*.

Altitude	680 km
Inclinação	98,1°
Velocidade	7 km/s
Duração da Órbita	98 min
Tipo de Órbita	Sol-síncrona
Resolução Espacial	Pancromática: 1m / Multiespectral: 4m
Bandas Espectrais	<div> <div>Pan</div> <div>0.45-0.90 μm</div> </div> <div> <div>Azul</div> <div>0.45-0.52 μm</div> </div> <div> <div>Verde</div> <div>0.52-0.60 μm</div> </div> <div> <div>Vermelho</div> <div>0.63-0.69 μm</div> </div> <div> <div>Infravermelho próximo</div> <div>0.76-0.90 μm</div> </div>
Resolução Radiométrica	11 bits (2048 níveis de cinza)
Imageamento	cenos de 13km x 13km
Período de Revisita	2.9 dias no modo Pancromático 1.5 dia no modo Multiespectral Esses valores valem para latitude de +/- 40°. A frequência de revisita para latitudes maiores será menor, e maior para as latitudes perto do Equador.

O satélite *IKONOS II* possui varredura tipo *push-broom*, ou seja, não tem espelho oscilatório ou rotativo e, por isso, não incorre no erro de não-linearidade (ou falha de construção) dos sensores. O sensor pode obter imagens estereoscópicas por mudança de aspecto orbital, ou seja, mudando o ângulo de imageamento em relação ao nadir. O *IKONOS II* pode obter imagens na forma nadiral e também em visadas inclinadas no sentido perpendicular a sua órbita, daí sua capacidade estereoscópica. Porém, ao fazê-lo, incorrerá na distorção panorâmica.

Conforme se pôde observar na tabela 1, as imagens obtidas pelo sensor *IKONOS II* possuem resolução espacial de 4m para as bandas multiespectrais, sendo 3 da faixa do visível e 1 do infravermelho próximo, e uma banda pancromática com

resolução espacial de 1m. A resolução radiométrica das imagens do sensor *IKONOS II* de 11 bits (2048 níveis de cinza) possui a vantagem do aumento do poder de contraste e de discriminação das imagens, inclusive nas áreas de sombra.

2.2 PROCESSAMENTO DIGITAL DE IMAGENS

De acordo com ALBUQUERQUE e ALBUQUERQUE (2000, p.10-22), o "Processamento de Imagens" se origina do Processamento de Sinais. Os sinais, como as imagens, são um suporte físico que contêm uma determinada *informação*, e esta pode estar associada a uma medida. Processar uma imagem consiste em transformá-la sucessivamente com o objetivo de extrair mais facilmente a *informação* nela presente.

A análise quantitativa e a interpretação de imagens representam pontos de apoio em diversas disciplinas científicas. Na realidade a diversidade de aplicações do processamento de imagens, está associada diretamente a análise da *informação* (ALBUQUERQUE e ALBUQUERQUE, 2000, p.10-22).

A área de processamento de imagens vêm apresentando um desenvolvimento expressivo nas últimas décadas. O nível de interesse vem aumentando na área acadêmica, onde os assuntos tais como morfologia matemática, redes neurais e compressão de imagens são objetos de pesquisas, teses e dissertações.

Para o processamento de imagens, algoritmos de computador são desenvolvidos e empregados. Estes algoritmos têm por finalidade alterar os valores digitais contidos nos *pixels* para facilitar a interpretação visual, ou deduzir informações a respeito dos objetos presentes na imagem a partir das variações espectrais registradas na mesma. Quando o processamento é feito *pixel a pixel*, apenas o valor do contador digital de cada elemento da imagem é considerado. No entanto, existem também algoritmos que analisam regiões, conjuntos de pixels com propriedades similares. Para isto, conceitos como "*vizinhança*" e "*conectividade*" são necessários para definir a região de interesse.

2.2.1 PROPRIEDADES DE UMA IMAGEM DIGITAL

Uma imagem digital $f(x,y)$ pode ser vista como uma matriz cujas linhas e colunas identificam um ponto na imagem. Da relação espacial entre os elementos desta matriz podem ser derivados os conceitos de "vizinhança", "conectividade" e "regiões", definidas a seguir.

a) Vizinhança

Um pixel p , com coordenadas (x, y) , possui 4 vizinhos horizontais e verticais, onde suas coordenadas são $(x+1, y)$, $(x-1, y)$, $(x, y+1)$, $(x, y-1)$. Estes pixels formam então a denominada *vizinhança-de-4* do pixel p , que é representado por $N_4(p)$ (SONKA *et al.*, 1998, p. 27).

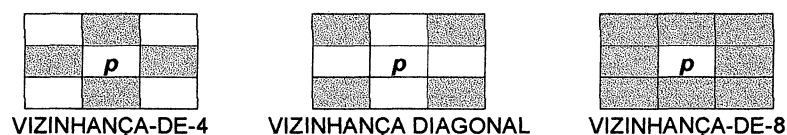
Os 4 vizinhos diagonais de p possuem como coordenadas: $(x+1, y+1)$, $(x+1, y-1)$, $(x-1, y+1)$, $(x-1, y-1)$, e estes serão denotados por $N_d(p)$.

A *vizinhança-de-8* de p é definida por:

$$N_8(p) = N_4(p) \cup N_d(p) \quad (6)$$

As *vizinhança-de-4*, *vizinhança diagonal* e *vizinhança-de-8* podem ser observadas através da Figura 4.

FIGURA 4 - CONCEITOS DE VIZINHANÇA-DE-4, VIZINHANÇA DIAGONAL E VIZINHANÇA-DE-8.



FONTE: MARQUES E VIEIRA, 1999, P. 35.

b) Conectividade

A conectividade entre pixels é um importante conceito utilizado no estabelecimento das bordas de objetos e componentes de regiões em uma imagem. De acordo com SONKA *et al.* (1998, p.28), para se estabelecer se dois pixels estão

conectados, é necessário determinar se eles são de alguma forma adjacentes (por exemplo, se são vizinhos-de-4) e seus níveis de cinza satisfazem em certo critério de similaridade. Por exemplo, em uma imagem binária, onde os pixels apenas podem assumir valores 0 e 1, dois pixels podem ser vizinhos-de-4, mas somente serão considerados 4-conectados se possuírem o mesmo valor.

O conceito de conectividade também pode ser definido para uma faixa de níveis de cinza. Seja V o conjunto de valores de níveis de cinza utilizados para se definir a conectividade. Por exemplo, em uma imagem binária, $V = \{1\}$ para a conexão de pixels com valor 1. Em uma imagem de múltiplos tons de cinza, para a conexão de pixels com valores de intensidade na faixa de 32 a 64, $V = \{32, 33, \dots, 63, 64\}$. Sabendo-se dos conceitos de vizinhança e dado o conjunto V , podem-se definir os critérios de conectividade (MARQUES e VIEIRA, 1999, p.35; GONZALES e WOODS, 2000, p.27-29):

b1. Conectividade-de-4: dois *pixels* p e q com valores de nível de cinza contidos em V , são conectados-de-4 se q está no conjunto $N_4(p)$.

b2. Conectividade-de-8: dois *pixels* p e q com valores de nível de cinza contidos em V , são conectados-de-8 se q está no conjunto $N_8(p)$.

b3. Conectividade-de- m (conectividade mista): dois *pixels* p e q com valores de nível de cinza contidos em V , são conectados-de- m se:

- (i) q está em $N_4(p)$ ou
- (ii) q está em $N_8(p)$ e o conjunto $N_4(p) \cap N_4(q)$ for vazio.

c) Regiões

De acordo com MASCARENHAS e VELASCO (1989, p. 8.11), “região é um conjunto de pontos conectados, ou seja, de qualquer ponto da mesma pode-se chegar a qualquer outro ponto por um caminho inteiramente contido na região”. As regiões que se deseja detectar, geralmente são regiões "homogêneas", ou seja, apresentam alguma propriedade local aproximadamente constante em toda sua extensão. Algumas

propriedades locais comumente utilizadas são nível de cinza, textura e cor.

O interesse em métodos de processamento de imagens decorre das duas principais áreas de aplicação, que são a melhoria de informação visual para a interpretação humana e a análise automática de informações extraídas de uma cena (MARQUES e VIEIRA, 1999, p.1-2; GONZALES e WOODS, 2000, p.1-4). As ferramentas do processamento digital podem ser usadas para identificar regiões conexas em uma imagem temática (método dos componentes conexos), para pós-processar imagens temáticas com o intuito de melhor caracterizar as regiões (por exemplo, usando a morfologia matemática) ou para obter descritores espaciais para as regiões.

2.3 RECONHECIMENTO DE PADRÕES

A exigência crescente de informatização e automação de tarefas humanas repetitivas e/ou cansativas dentro do processo de análise de imagens levam a desenvolver ferramentas específicas para atender projetos cujo conhecimento inicial é incompleto e onde um certo grau de inteligência é importante. O reconhecimento de um padrão inscreve-se neste tipo de tarefa.

A abrangência do reconhecimento de padrões está na sua efetiva realização para as diversas áreas de pesquisas. A área abrange desde a detecção de padrões à escolha mais simples entre dois objetos, como a complexa realização da aprendizagem. Na prática computacional, pode-se enumerar algumas áreas onde vem sendo usado o reconhecimento de padrões (GARCIA, 1982): análise, segmentação, classificação de imagens; reconhecimento de faces; identificação de impressões digitais; reconhecimento de caracteres; diagnóstico médico.

2.3.1 CLASSIFICAÇÃO DE IMAGENS

As técnicas de classificação e interpretação de imagens são utilizadas em diversos projetos, tais como: monitoramento das mudanças dos níveis d'água de lagos e reservatórios, avaliação de doenças de plantas, localização de novas rotas de estradas, identificação e avaliação de mudanças de uso do solo. A interpretação de imagens pode ser definida como o processo de reconhecimento de objetos em fotografias aéreas ou em imagens de sensoriamento remoto, analisando suas características espectrais, sua forma, tamanho, textura, sombras e as associações existentes entre estes elementos (AVERY e BERLIN, 1992, p.51). Uma das maneiras de atingir a interpretação automática de uma imagem é através de processos de classificação digital, onde pixels, ou partes da imagem, são analisados e o tipo de objeto ou cobertura mais provável é determinado em função de suas propriedades espectrais, espaciais e/ou de textura.

A classificação de uma imagem digital é uma tarefa complexa. Não existe uma única solução geral para este problema, pois cada imagem possui propriedades diferentes em função de vários fatores, como sua escala, o tipo de objetos presentes na cena, o ângulo de visada, a iluminação, etc.. Embora muitas destas dificuldades não sejam difíceis de serem superadas pelo cérebro humano, a análise de uma imagem usando computadores demanda a fixação dos parâmetros a serem analisados e um modelo que possa ser aplicado com sucesso. Como não existe um modelo geral, existem diferentes abordagens para a solução deste problema. Duas grandes linhas podem ser identificadas:

- abordagem *pixel a pixel*;
- abordagem *orientada a regiões, ou orientadas ao objeto*, que é seguida de dois processos:
 - segmentação;
 - classificação.

Na primeira, os elementos mínimos da imagem, os *pixels*, são analisados isoladamente. Cada um deles é considerado um elemento independente e uma decisão é tomada a respeito da interpretação mais apropriada em função de suas únicas propriedades, as suas informações espectrais. Este conceito foi muito aplicado na classificação de imagens de satélite com baixa ou média resolução espacial, pois nestas imagens a discriminação de áreas ocupadas por diferentes coberturas no terreno podia ser feita analisando apenas as diferenças espectrais. A abordagem "*pixel a pixel*" é eficiente em termos computacionais, pois cada elemento da imagem é processado usando o mesmo algoritmo de maneira sequencial. Grande parte dos métodos de classificação de imagens de sensoriamento remoto se baseiam neste princípio. Pode-se citar, por exemplo, a classificação pelo método da máxima verossimilhança disponível em programas como o *ENVI* e *ERDAS*.

Com o aumento da resolução espacial das imagens, os objetos presentes na cena podem ser melhor diferenciados em uma imagem e suas características espectrais não são suficientes para uma adequada discriminação. Torna-se então necessário integrar outras informações, como forma ou textura de regiões da imagem. A análise de regiões é uma opção à análise *pixel a pixel* que demanda um maior esforço computacional, visto que inicialmente é necessário caracterizar as regiões e descrevê-las de maneira adequada. Este tipo de análise é chamada "*orientada a objeto*" ou "*orientada a regiões*".

Na abordagem orientada a regiões, a imagem deve ser inicialmente dividida em regiões coerentes, para depois classificar estas regiões segundo suas características espectrais e espaciais.

De acordo com NISHIDA (1998), os classificadores tradicionais *pixel-a-pixel* apenas utilizam as informações espectrais, sendo que este fato mostra-se como uma limitação na utilização dos mesmos. Com a inserção dos novos sensores de alta resolução espacial, onde maiores detalhes da paisagem urbana podem ser identificados, os métodos de classificação orientados ao objeto passam a ter maior

importância, já que os mesmos consideram além dos atributos espectrais as relações espaciais entre os pixels.

De acordo com KUO; CHOU e LEE (2001), nos últimos anos vêm surgindo sensores com alta resolução espacial, como o *IKONOS II* com uma resolução espacial de 1m, e muitos outros sensores estão sendo desenvolvidos como o *QUICKBIRD* com resolução espacial de 61cm. Muitos detalhes como construções, estradas, árvores como também outros componentes que podem ser claramente identificados nestas imagens de alta resolução espacial. Com isto, os métodos de classificação baseados em *regiões* ou *objetos* passam a ser mais adequados.

SCHNEIDER e STEINWENDNER (1999) utilizaram a segmentação e a classificação de imagens de satélite para a obtenção de mapas de uso da terra e segundo os autores, o bom resultado da classificação depende do resultado da segmentação com a inserção de parâmetros de forma e textura.

MOREIRA e ASSAD (1998) utilizaram a segmentação e classificação na obtenção de mapa de uso da terra na Microbacia do córrego Lamarão no Distrito Federal, onde para tanto dividiram as imagens Landsat 5 em regiões e extraíram os atributos espectrais e geométricos que foram utilizados para posterior classificação destas regiões. De acordo com os autores, a utilização das técnicas de segmentação de imagens e classificação supervisionada por regiões mostraram-se eficientes na separabilidade das regiões, sendo que muitas bordas dos segmentos coincidiram com as definidas pela análise visual.

A diferença principal entre estas duas abordagens é o objeto da classificação, que pode ser o *pixel* ou uma *região*, resultado da segmentação da imagem.

Segundo CHUVIECO (1990, p.325), o processo de classificação de imagens é de grande importância no processamento digital de imagens, e este consiste na associação de cada elemento da mesma a um "rótulo" que descreve um objeto da superfície terrestre gerando uma nova imagem, onde é associada uma cor para cada classe e todos os elementos pertencentes a esta classe possuem a mesma cor. Quando

este tipo de operação é efetuado para todos os elementos da imagem, o resultado é um mapa temático, mostrando a distribuição geográfica das classes que compõe a cena classificada. Aqui se entende como "*elemento*" um *pixel* ou *região* da imagem.

De acordo com LILLESAND e KIEFER (1994), conforme o processo de classificação empregado, os classificadores podem ser divididos em classificadores *pixel a pixel* e classificadores *por regiões*.

Os classificadores *pixel a pixel* utilizam a informação espectral de cada pixel na busca por regiões homogêneas.

A classificação também pode ser aplicada a regiões previamente identificadas. Os classificadores *por regiões* utilizam além da informação espectral de cada pixel, a informação espacial que envolve a relação entre os pixels e seus vizinhos. Estes classificadores procuram simular o comportamento de um foto-intérprete, ao reconhecer áreas homogêneas de imagens, baseados nas propriedades espectrais e espaciais de imagens (LILLESAND e KIEFER, 1994).

A classificação é expressa pelas seguintes fases:

- definição das classes (fase de treinamento do classificador);
- distribuição dos pixels da imagem em uma dessas classes (fase de destino), e;
- verificação dos resultados (acurácia da classificação).

2.3.1.1 Fase de Treinamento do Classificador

Esta fase define cada uma das classes que se pretende discriminar, levando em consideração sua própria variabilidade na área em estudo. Este processo define a seleção de amostras na imagem, que representa adequadamente as categorias de interesse. A partir destas amostras, pode-se calcular os valores das variáveis que descrevem/caracterizam a classe. No caso de uma classificação *pixel a pixel*, serão utilizados os níveis de cinza médios e a variância. No caso da classificação *por regiões* serão estimados os valores médios e a dispersão de variáveis espectrais e/ou espaciais, como a forma dos segmentos. Pode-se considerar a fase de treinamento das amostras

como a coluna vertebral da classificação numérica, pois os resultados da mesma estão diretamente relacionados com a definição prévia das classes.

Segundo LILLESAND e KIEFER (1994), teoricamente, o número mínimo de elementos que o conjunto de treinamento deve possuir é $n + 1$, onde n é a dimensionalidade das variáveis. Na prática, segundo CHUVIECO (1990, p.329), o número mínimo de elementos varia de $10n$ a $100n$. Isto se deve ao fato os valores estimados da média e das matrizes de covariância serem mais precisos à medida que o número de pixels do conjunto aumenta.

2.3.1.2 Fase de Destino

Realizada a seleção e estudo das categorias que entraram na classificação, inicia-se a fase que trata da distribuição de cada elemento da imagem (pixel ou região) a uma das classes previamente selecionadas. Em Sensoriamento Remoto, vários métodos são utilizados, tais como: distância mínima, método do paralelepípedo e o método da máxima verossimilhança (ou geral Gaussiana).

No método do paralelepípedo, os *pixels* classificados dentro de uma classe (C) são aqueles com *níveis de cinza* que encontram-se dentro de um certo limite, determinado pelo cálculo do valor médio dos *níveis de cinza* para cada classe (M_C) e seu desvio padrão (DP_C). Assim, os *pixels* que estiverem dentro do limite que vai de $(M_C - DP_C)$ até $(M_C + DP_C)$, são classificados como pertencentes à classe C .

O método da distância mínima considera o *pixels* resultante da média dos *níveis de cinza* para cada classe em todas as bandas. São calculadas as distâncias Euclidianas entre determinado *pixel* (x) apresentado ao sistema e o *pixels* de cada uma das classes. Como o próprio nome sugere, o *pixels* apresentado é classificado na classe para a qual a distância Euclidiana resultou em um valor mínimo.

O último método, mais comumente utilizado é o da máxima verossimilhança Gaussiana. Este método pressupõe que os elementos de cada classe seguem uma distribuição normal multivariada, e o método avalia as probabilidades que um

determinado elemento tem de pertencer às classes identificadas e o designa àquela cuja probabilidade é maior (NISHIDA, 1998).

2.3.1.3 Avaliação da Classificação

Toda classificação leva consigo uma certa margem de erro, isto em função da qualidade dos dados e do método empregado. Por este motivo há a necessidade da aplicação de algum método de verificação, que permita quantificar os possíveis erros, validando assim a qualidade do produto final.

A precisão pode ser medida tomando-se como base a "verdade de campo", conhecida a partir de trabalhos realizados em campo, da interpretação de fotografias aéreas ou ainda de informações geográficas a respeito do terreno. Então são comparadas a precisão dos mapas que contenham a verdade terrestre e a dos obtidos pela classificação. É necessária a identificação das classes em ambos os mapas, para a formação posterior de uma "matriz de confusão".

A matriz de confusão tem tamanho $m \times m$, sendo m o número de classes já definidas. Sua diagonal compreende o número de elementos corretamente classificados, suas linhas representam as classes da verdade de campo e suas colunas as classes do mapa obtido através da classificação.

Uma das principais características destas matrizes é o fato de resumirem os erros de inclusão (elementos que não pertencem à classe e foram a ela atribuídos) e os erros de omissão (elementos que pertencem à classe em questão e não foram a ela atribuídos). A partir da matriz de confusão pode-se calcular um coeficiente de precisão da classificação total. O mais empregado para classificações temáticas por representar toda a matriz de confusão, levando em consideração os erros de inclusão e omissão, é o coeficiente Kappa (κ) (GONG e HOWARTH, 1992, p. 423-437) dado por:

$$k = \frac{N \cdot \sum_{i=1}^M X_{ii} - \sum_{i=1}^M (X_{i+} \cdot X_{+i})}{N^2 - \sum_{i=1}^M (X_{i+} \cdot X_{+i})} \quad (7)$$

onde :

N – número total de pixels da amostra

M – total de classes

$\sum_{i=1}^M X_{ii}$ – soma total pixels classificados corretamente

X_{i+} – soma pixels das linhas (erro de omissão)

X_{+i} – soma pixels das colunas (erro de inclusão)

O coeficiente *Kappa* pode apresentar valores entre -1 e +1, e o mesmo pode mostrar o desempenho da classificação de acordo com a seguinte Tabela 2:

TABELA 2 - COEFICIENTE KAPPA E O DESEMPENHO DA CLASSIFICAÇÃO

Kappa	Desempenho da Classificação
< 0	Péssimo
$0 < \kappa < 0,2$	Mau
$0,2 < \kappa < 0,4$	Razoável
$0,4 < \kappa < 0,6$	Bom
$0,6 < \kappa < 0,8$	Muito Bom
$0,8 < \kappa < 1,0$	Excelente

FONTE: WWW.DPI.INPE/BR/~LEILA/SER437/PDI6-SEG-CLASS.PDF EM 16/10/2001.

2.3.2 SEGMENTAÇÃO DE IMAGENS

ALBUQUERQUE e ALBUQUERQUE (2000) consideram que o processo de segmentação de imagens consiste na primeira etapa de processamento da imagem, dentro de uma análise do ponto de vista da *informação* nela presente. Existem diversas técnicas de segmentação de imagens, mas não existe um método único que seja capaz de segmentar todos os tipos de imagem.

De acordo com GONZALES e WOODS (2000, p.295), a segmentação é um processo em que uma imagem é subdividida em partes ou objetos constituintes. O

nível até o qual essa subdivisão deve ser realizada depende do problema a ser resolvido. Ou seja, a segmentação deverá ser finalizada quando os objetos de interesse na aplicação tiverem sido isolados.

Os algoritmos de segmentação de imagens são muito específicos, ou seja, desenvolvidos para determinada aplicação. O algoritmo que detecta com grande precisão um tumor não obteria o mesmo resultado ao tentar detectar um defeito de uma peça numa linha de montagem, portanto, não existe um modelo geral de segmentação.

2.3.2.1 Técnicas de Segmentação

Os algoritmos de segmentação são geralmente baseados em uma das seguintes propriedades básicas de valores de cinza: descontinuidade e similaridade. Na primeira categoria a abordagem é particionar a imagem baseando-se em mudanças bruscas nos níveis de cinza. As principais áreas de interesse nessa categoria são a detecção de pontos isolados e detecção de linhas e bordas na imagem. Na segunda abordagem, *pixels* adjacentes, com propriedades similares são agrupados para formar regiões uniformes. Os principais algoritmos da segunda categoria baseiam-se em limiarização, crescimento de regiões, divisão e fusão de regiões (GONZALES e WOODS, 2000, p.295-344).

2.3.2.1.1 Detecção de Bordas

As técnicas baseadas na detecção de bordas têm como objetivo identificar descontinuidades de intensidade, que correspondem às fronteiras dos objetos na imagem.

Na prática, o conjunto de pixels resultante do processo de detecção de bordas, raramente caracteriza completamente uma fronteira devido a efeitos que introduzem descontinuidade de intensidade, como iluminação não uniforme, causando então quebras na fronteira. Portanto, os algoritmos de detecção de bordas são geralmente

seguidos de procedimentos de ligação de bordas, como o de "perseguição de contorno".

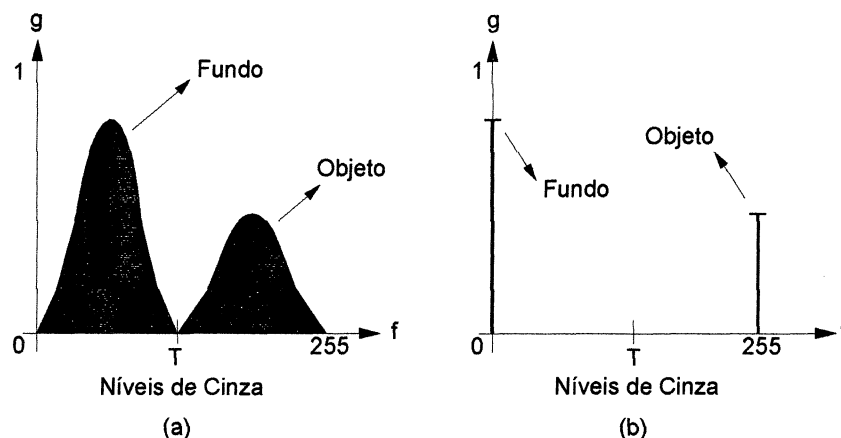
A ligação de borda é feita sobre uma imagem resultante da extração de bordas (por exemplo uma imagem gradiente que saliente as bordas, ou contornos, na imagem). O processo de perseguição de contorno se inicia quando se detecta um pixel com valor superior ao limiar estabelecido (T). A partir deste *pixel*, observa-se a vizinhança para identificar o o *pixel* vizinho com maior valor de nível digital, que é considerado continuação da borda. Pode ocorrer que exista mais de uma alternativa. Neste caso, das possíveis continuações escolhe-se uma aleatoriamente. Tomando este *pixel* vizinho como novo ponto central da vizinhança, repete-se o processo e entre os pontos examinados são excluídos aqueles que pertencem a algum contorno. Desse processo gera-se uma imagem binária com os valores de 1 referentes às bordas e 0, a regiões de não borda (MARQUES e VIEIRA, 1999, p.51).

2.3.2.1.2 Limiarização (*Thresholding*)

De acordo com MARQUES e VIEIRA (1999, p.96), o princípio da limiarização consiste em separar as regiões de uma imagem quando esta apresenta duas classes (o fundo e o objeto). As técnicas de limiarização são as mais simples, pois não levam em conta o contexto, ou seja, a vizinhança dos pixels não é analisada. A técnica consiste em definir um limiar de separação, a partir do qual a imagem será segmentada. Uma aplicação básica onde esta técnica pode ser utilizada é na binarização de imagens, que transforma uma imagem com diversos tons de cinza em uma imagem com apenas dois tons, por exemplo preto e branco. Esta técnica pode ser estendida à estimativa de vários limiares (multilimiarização).

A forma mais simples e geral de limiarização consiste na bipartição do histograma, convertendo os pixels cujo tom de cinza é maior ou igual a um certo valor de limiar (T) em brancos e os demais em pretos (Figura 5).

FIGURA 5 - LIMIAÇÃO DE UMA IMAGEM MONOCROMÁTICA UTILIZANDO LIMIAIR T : (a) HISTOGRAMA ORIGINAL, (b) HISTOGRAMA DA IMAGEM BINARIZADA.



FONTE: MARQUES E VIEIRA, 1999, P.96.

Matematicamente, a operação de limiarização pode ser descrita como uma técnica de processamento de imagens na qual uma imagem de entrada $f(x,y)$ de N níveis de cinza produz uma imagem de saída $g(x,y)$, chamada limiarizada, cujo número de níveis de cinzas é menor que N . Normalmente, $g(x,y)$ apresenta 2 níveis de cinza, sendo:

$$g(x,y) = 1 \text{ se } f(x,y) \geq T \quad (10)$$

$$g(x,y) = 0 \text{ se } f(x,y) < T \quad (11)$$

onde os *pixels* rotulados com 1 correspondem aos objetos e os *pixels* etiquetados com 0 correspondem ao fundo e T é um valor de nível de cinza pré-definido, ao qual denomina-se limiar (SONKA *et al.*, 1998, p.124).

2.3.2.1.3 Segmentação Orientada a Regiões

As regiões que se deseja detectar, em geral, são regiões “homogêneas”, ou seja, apresentam alguma propriedade local aproximadamente constante em toda sua extensão. Algumas propriedades locais usadas comumente são o nível de cinza, a textura e a cor. Seja R a imagem completa, pode-se imaginar a segmentação como o processo de particionar R em n regiões R_1, R_2, \dots, R_n , tal que:

$$(a) \quad \bigcup_{i=1}^n R_i = R \quad (12)$$

$$(b) \quad R_i \text{ é uma região conexa, } i = 1, 2, \dots, n$$

$$(c) \quad R_i \cap R_j = \emptyset \text{ para todo } i \text{ e } j, i \neq j \quad (13)$$

$$(d) \quad P(R_i) = \text{VERDADEIRO para } i = 1, 2, \dots, n, \text{ e}$$

$$(e) \quad P(R_i \cup R_j) = \text{FALSO para } i \neq j \quad (14)$$

em que $P(R_i)$ é um predicado lógico sobre os pontos do conjunto R_i e \emptyset é o conjunto vazio.

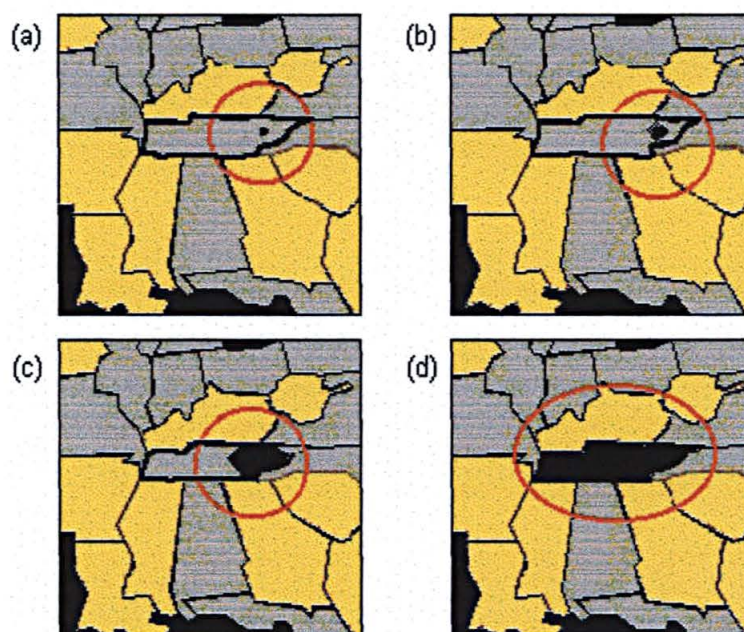
A condição (a) indica que a segmentação deve ser completa, ou seja, cada *pixel* deve pertencer a uma região. A condição (b) requer que os *pixels* em uma região sejam conexos. A condição (c) indica que as regiões devem ser disjuntas. A condição (d) trata das propriedades que devem ser satisfeitas pelos *pixels* em uma região segmentada, por exemplo, $P(R_i) = \text{VERDADEIRO}$ se todos os *pixels* em R_i possuírem a mesma intensidade. Finalmente, a condição (e) indica que as regiões R_i e R_j são diferentes no sentido do predicado P (GONZALES e WOODS, 2000, p.326).

2.3.2.1.3.1 Crescimento de Regiões por Agregação de Pixels

O princípio desta técnica consiste em agregar *pixels* similares que ocorrem em posições adjacentes. Começa-se com um conjunto de pontos "semente" e, a partir deles, crescem as regiões anexando a cada ponto semente aqueles *pixels* que possuam propriedades similares. A propriedade a ser considerada pode ser o nível de cinza do *pixel*, sua cor ou textura. O resultado é um conjunto de regiões uniformes, que podem ser entendidas como novos elementos da imagem e podem ser analisadas de acordo aos valores digitais de seus *pixels*, como também de acordo a sua forma.

A Figura 6 mostra uma seção de um mapa contendo um único ponto semente. Percebe-se que, embora *pixels* nas regiões vizinhas satisfaçam o critério de nível de cinza, o crescimento parou devido a esses *pixels* não satisfazerem o critério de conectividade, causado pela separação criada pela borda escura em torno da região crescida.

FIGURA 6 - (a) IMAGEM ORIGINAL USANDO UM PONTO SEMENTE; (b) ESTÁGIO PRIMÁRIO DE CRESCIMENTO DE UMA REGIÃO; (c) ESTÁGIO INTERMEDIÁRIO DE CRESCIMENTO DE UMA REGIÃO; (d) ESTÁGIO FINAL.



FONTE: ADAPTADO DE GONZALES E WOODS, 2000, p.328.

2.3.3 REPRESENTAÇÃO E DESCRIÇÃO DE ELEMENTOS DA IMAGEM

Reconhecimento é a parte do processamento para a classificação de objetos a partir de informações encontradas na imagem, geralmente tendo como apoio uma base de conhecimento previamente estabelecida (GUYON, 1991, p.215-259). Esta fase é normalmente aplicada após uma fase de segmentação e a parametrização da imagem. A fase de parametrização identifica e calcula parâmetros nos objetos segmentados. Um exemplo de parâmetros pertinentes a um objeto é o seu perímetro ou sua área (FU; MUI, 1980, p.3-16 *apud* ALBUQUERQUE e ALBUQUERQUE, 2000).

Para serem submetidos ao processo de classificação, os agrupamentos resultantes de *pixels* segmentados ou resultantes de classificação devem ser representados e descritos através de parâmetros que possam ser processados usando computadores. A representação de uma região envolve duas escolhas:

- através da representação da região em termos de suas características externas (sua fronteira), ou
- em termos de suas características internas (os *pixels* que compõem a região).

Geralmente, uma representação externa é escolhida quando a atenção primária estiver voltada para as características de forma. Na representação interna, são voltadas às propriedades como cor e textura (GONZALES e WOODS, 2000, p.345).

2.3.3.1 Textura

Uma textura pode ser classificada como artificial ou natural. Texturas artificiais consistem num arranjo de símbolos colocados sobre um fundo neutro. Tais símbolos podem ser segmentos, por exemplo. Texturas naturais são imagens de cenas naturais contendo um arranjo semi-repetitivo de *pixels*. Como exemplos estão as imagens de tecidos, grama, areia, etc.

As três abordagens principais usadas em processamento de imagens para a descrição de texturas são a estatística, a estrutural e a espectral. As abordagens estatísticas levam a caracterizações de textura como suave, áspera, granular. As técnicas estruturais tratam arranjos de primitivas de imagem, como a descrição da textura baseada em linhas paralelas regularmente espaçadas. As técnicas espectrais baseiam-se em propriedades do espectro de *Fourier* (GONZALES e WOODS, 2000, p.361-367).

Inúmeras técnicas de análise de textura têm sido desenvolvidas na literatura de processamento de imagens. LIBERMAN (1997) por exemplo, utilizou a textura no auxílio à classificação de imagens digitais com o auxílio de redes neurais; GALO (2000) e MARTINS; GUIMARÃES e FONSECA (2002) utilizaram características de textura, aplicando a metodologia das redes neurais na classificação de imagens de sensoriamento remoto. DAHMER (1998) utilizou o auxílio de parâmetros de textura na segmentação de imagens Ecocardiográficas. Uma definição precisa de textura parece não ter sido ainda proposta, embora seja poderosa a capacidade da visão humana em sua discriminação (MASCARENHAS e VELASCO, 1989, p.9.27).

Conforme SONKA *et al.* (1998, p.647), a textura refere-se a propriedades referentes às superfícies ou estruturas dos objetos, tais como suavidade, rugosidade e regularidade.

2.3.3.2 Forma

A forma pode ser considerada como uma expressão de contorno dos objetos. Objetos formados por culturas ou formados pelas construções observadas na paisagem urbana, normalmente tendem a apresentar formas regulares, enquanto que feições naturais tendem a apresentar formas irregulares (GARCIA, 1982; AVERY e BERLIN, 1992, p.51-58).

A análise de forma e de tamanho é uma poderosa ferramenta para a

discriminação de objetos que possuem a mesma aparência espectral. Assim como a análise espectral baseia-se nos valores digitais das bandas, é necessário dispor de um parâmetro de forma e tamanho para descrever a forma dos objetos. Para analisar a forma é necessário conhecer os limites dos objetos nas imagens, algo que a princípio é desconhecido. Por isto, os algoritmos de análise de forma partem de uma imagem temática, onde os objetos se encontrem mais definidos. A imagem temática pode ser o resultado de uma classificação supervisionada, ou a segmentação da imagem.

Um exemplo de classificação pela análise da forma pode ser encontrado em ALBUQUERQUE; ALBUQUERQUE (2000). Em seu trabalho, eles adotaram os parâmetros coeficiente de compacidade (cc) e a relação entre os eixos menor (b) e maior (a) da *ellipse* que mais se aproxima do objeto para classificar duas classes, sendo uma de canetas e a outra de alicates, e puderam concluir que apenas um dos parâmetros já seria suficiente para a referida classificação.

Outra maneira de analisar a forma é a partir da Teoria dos Fractais. DE COLA (1989) utilizou esta técnica para a obtenção das dimensões fractais para alguns objetos encontrados em uma imagem Landsat de Vermont, e pôde concluir que a teoria dos fractais é uma forma de descrever a complexidade dos contornos dos padrões que não são facilmente descritos pela geometria euclidiana, porém possui a desvantagem de ser variante à rotação da imagem quando aplicada a imagens digitais (CENTENO *et al*, 2001).

O parâmetro de forma ideal não deve depender de variações de escala e rotação da imagem. Entre os parâmetros utilizados como descritores da variação espacial do objeto encontram-se (BÄSSMAN e KREYSS, 1998; SONKA *et al*, 1998, p.229-274):

- área (A);
- perímetro (P);
- coeficiente de compacidade (cc);
- assinaturas (momentos);
- relação entre os eixos menor e maior da *ellipse* que melhor se

aproxima do objeto (excentricidade);

- código da cadeia;
- segmentos de fronteiras (deficiência convexa).

Há várias outras maneiras de representar e descrever a variação da forma dos objetos, como através dos descritores de *Fourier*, fineza, princípio de Pavlidis, entre outras (GONZALES e WOODS, 2000, p.345-359; FACON, 1993, p.148-171).

2.3.3.2.1 Área

Uma das propriedades mais simples de uma região, a área, é definida como o número de *pixels* que compõem por uma região. O cálculo da área pode ser realizado através de algoritmos simples, onde o valor encontrado é resultado da soma de todos os *pixels* pertencentes ao objeto analisado (MASCARENHAS e VELASCO, 1989, p.9.18; SONKA *et al.*, 1998, p.254). Este descritor é variante a mudanças de escala (FACON, 1993, p.151).

2.3.3.2.2 Perímetro

O perímetro assim como a área, é uma das características básicas de uma região. É facilmente calculado considerando o número de *pixels* conexos (em função do predicado de conectividade escolhido), que constituem o contorno da região. O perímetro é o tamanho de sua fronteira.

Este descritor é invariante à translação, porém geralmente deve ser aplicado a situações em que o tamanho e a orientação dos objetos não varie (GONZALES e WOODS, 2000, p.359). A variação quanto à rotação se deve ao fato de que a distância entre dois pontos localizados sobre os eixos horizontais ou verticais vale 1 e a distância para dois pontos sobre a duas diagonais vale $\sqrt{2}$.

2.3.3.2.3 Coeficiente de Compacidade

De acordo com BÄSSMAN e KREYSS (1998), o coeficiente de compacidade mede o grau de semelhança entre a região e uma circunferência de igual área. Normalmente os fatores de forma são definidos para variar entre 0 e 1, assumindo valor máximo para objetos com formas geométricas específicas. Este parâmetro descreve quanto à forma se aproxima de uma circunferência com a mesma área, por isto o valor do coeficiente de compacidade de uma circunferência é 1. Este valor cresce na medida em que o contorno do objeto se torna irregular, ou seja, menos compacto (SONKA *et al*, 1998, p.259). O coeficiente de compacidade (cc), é obtido através da equação (23):

$$cc = \frac{P^2}{4\pi A} \quad (15)$$

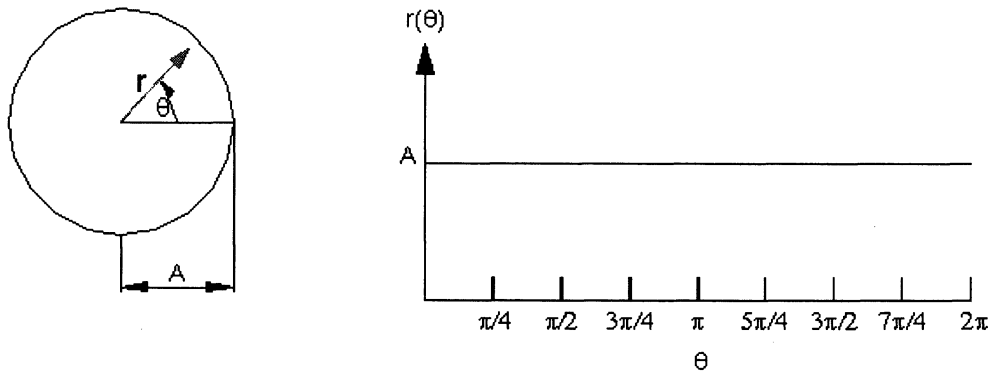
onde P corresponde ao valor do perímetro, e A ao valor da área do objeto analisado.

De acordo com FACON (1993, p.152), o coeficiente de compacidade (cc) permite, a partir de descritores simples como o perímetro e a área que não são invariantes no caso de mudança de tamanho do objeto, ter um descritor invariante à mudança de escala. A compacidade é também invariante à translação e à mudança de orientação.

2.3.3.2.4 Assinaturas

Uma assinatura é a representação unidimensional de uma fronteira, podendo se gerada através de um gráfico da distância da fronteira ao centróide em função do ângulo, Figura 7. A idéia para a utilização de uma assinatura é reduzir a representação do contorno a uma função unidimensional a qual permite uma descrição mais simples (GONZALES e WOODS, 2000, p.348; FACON, 1993, p.153).

FIGURA 7 – FRONTEIRA SIMPLES E SUA CORRESPONDENTE ASSINATURA.



FONTE: ADAPTADO DE GONZALES E WOODS, 2000.

A forma das assinaturas de uma fronteira pode ser descrita quantitativamente através dos momentos. HU (1962), definiu a partir dos momentos um conjunto de 7 momentos invariantes pela translação, rotação e mudanças de escala.

Conceitua-se por momentos invariantes o conjunto de funções invariantes à escala, translação e rotação, obtidas a partir dos momentos geométricos da imagem. Dada uma imagem digital, a equação dos momentos de ordem $(p + q)$, é representada por (FACON, 1993, p.153; VARELLA *et al.*, 2000; SONKA *et al.*, 1998, p.259):

$$m_{pq} = \sum_i \sum_j x_i^p y_j^q f(i, j) \quad (16)$$

para $p, q = 0, 1, 2, \dots$, e onde $f(i, j)$ representa o nível de cinza do ponto $P(i, j)$.

Os momentos permitem definir as propriedades de base de uma região utilizadas:

- área: essa propriedade comumente usada corresponde ao momento m_{00} . Esse momento representa o número de *pixels* constituindo a região:

$$m_{00} = \sum_i \sum_j f(i, j) \quad (17)$$

- centro de gravidade: definido da seguinte maneira:

$$\begin{aligned}x_g &= m_{10} / m_{00} \\y_g &= m_{01} / m_{00}\end{aligned}\tag{18}$$

A orientação v do grande eixo de inércia pode ser determinada a partir dos momentos centrados $\mu_{01}, \mu_{10}, \mu_{11}, \mu_{02}$ e μ_{20} :

$$\eta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^\gamma}, \quad \text{onde } \gamma = \frac{p+q}{2} + 1\tag{19}$$

onde μ_{pq} representam os momentos centrados de ordem $p + q$ cuja fórmula é dada por:

$$\mu_{pq} = \sum_i \sum_j (x_i - x_j)^p (y_i - y_j)^q f(x, j)\tag{20}$$

De acordo com SONKA (1998, p.259), este tipo de descrição é invariável pela translação e mudança de escala. Para eliminar o efeito de rotação, aplicam-se os 7 momentos normalizados, que são obtidos a partir dos momentos μ_{pq} e estes são invariantes a translação, rotação e mudança de escala, e são dados por:

$$\begin{aligned}\phi_1 &= \eta_{20} + \eta_{02}; \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4(\eta_{11})^2; \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2; \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2; \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2 + \\ &\quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]]; \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ &\quad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{03} + \eta_{21}); \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] - \\ &\quad (\eta_{30} - 3\eta_{12})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2].\end{aligned}\tag{21}$$

Como os valores de ϕ_1 a ϕ_7 são extremamente pequenos, é comum, ao final do cálculo dos momentos, aplicar-se a função logarítmica a fim de se evitar problemas de precisão numérica. Logo, as características extraídas dos objetos são definidas por:

$$\log|\phi_i|, i = 1, \dots, 6 \quad (22)$$

VARELLA *et al.* (2000), utilizaram 6 momentos invariantes no reconhecimento de caracteres com variações de escala, rotação e translação, e puderam concluir que para imagens digitais os resultados mostraram pequenas variações, especialmente, quando se efetuam rotações maiores que 45 graus. Os autores foram motivados à utilização deste descritor pelas publicações de DUDANI (1977, p.39-46) e KHOTANZAD (1988, p.625-632).

2.3.3.2.5 Relação Eixo Menor e Maior da Elipse

Este parâmetro descreve o grau de alongamento de uma região. Ele é calculado através da razão dos eixos menor e maior da elipse que mais se ajusta ao objeto ou elemento analisado (MASCARENHAS e VELASCO, 1998, p.9.22). Estes eixos podem ser obtidos através da análise por componentes principais, onde o cálculo destes eixos é obtido através dos vetores próprios da matriz variância covariância obtida usando os *pixels* constituindo a região como variáveis aleatórias.

No cálculo da transformação por componentes principais, cada *pixel* da imagem, ou de uma região, é tratado como um vetor bidimensional, e estes vetores são usados para computar o vetor médio e a matriz variância covariância do conjunto (região). A partir da determinação da variância dos *pixels* pertencentes à região, são calculados os autovetores que correspondem à orientação das 1ª e 2ª componentes, os quais são ortogonais, e autovalores que representam o comprimento dos eixos das componentes principais de uma imagem (CROSTA, 1992, p.140).

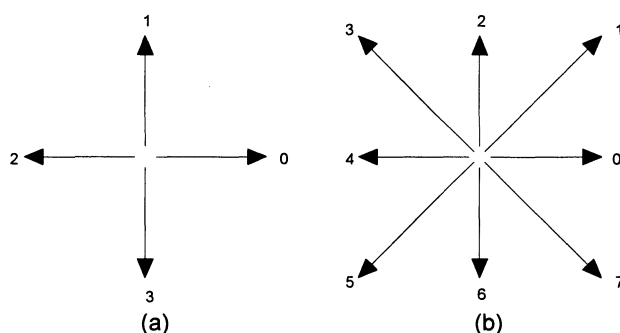
O objetivo desta transformação é alinhar um objeto bidimensional com o seu

autovetor principal, onde se pode então após este alinhamento, obter o comprimento dos eixos maior e menor do objeto analisado.

2.3.3.2.6 Código da Cadeia

O código da cadeia é usado na representação da fronteira por uma sequência conectada de segmentos de linha reta de determinado tamanho e direção. Essa representação baseia-se tipicamente na conectividade-de-4 ou na de-8 dos segmentos. A direção de cada segmento é codificada por um esquema de numeração como o mostrado na Figura 8 (GONZALES e WOODS, 2000, p.346; BÄHR,1991, p.121).

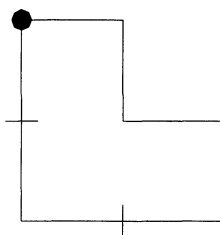
FIGURA 8 - DIREÇÕES DO (A) CÓDIGO DA CADEIA DIRECIONAL-DE-4 E (B) DIRECIONAL-DE-8.



FONTE: GONZALES E WOODS, 2000

Um exemplo simples para a representação deste descritor pode ser observado através da Figura 9, o início da descrição do código é no canto superior esquerdo da forma.

FIGURA 9 - REPRESENTAÇÃO DA FORMA ATRAVÉS DO CÓDIGO DA CADEIA USANDO 4 DIREÇÕES.



Código-da-cadeia: 0 3 0 3 2 2 1 1

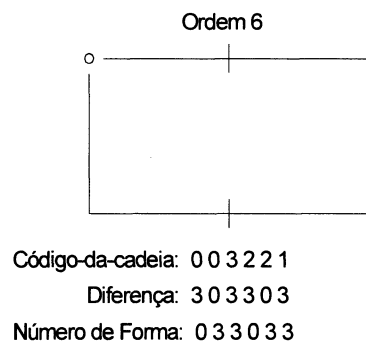
FONTE: GONZALES E WOODS, 2000.

Como o código da cadeia depende do ponto inicial e da rotação da região, há a necessidade da sua normalização, e a normalização realizada quanto à rotação é feita utilizando a primeira diferença do código no lugar do próprio código. A diferença é obtida contando-se (no sentido anti-horário) o número de direções que separam cada par de códigos adjacentes. Por exemplo, a primeira diferença do código da cadeia direcional-4 de 10103322 é 3133030.

A normalização em relação ao ponto de início do código é realizada através do seguinte procedimento: o código é gerado começando de um ponto arbitrário como uma seqüência circular de números, e então redefinido o ponto inicial de maneira que a seqüência de números normalizado quanto à orientação forme um inteiro de menor magnitude. Para o exemplo anterior, a codificação ficaria 0303133.

A Figura 10 mostra a representação da forma de ordem 6, juntamente com o código da cadeia e suas normalizações.

FIGURA 10 - REPRESENTAÇÃO DO CÓDIGO DA CADEIA NORMALIZADO.



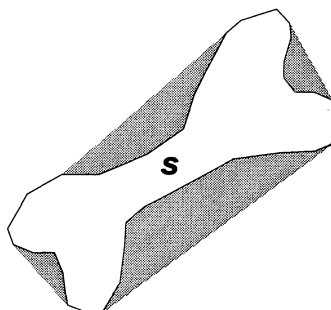
FONTE: ADAPTADO DE GONZALES E WOODS, 2000.

2.3.3.2.7 O fecho convexo

O *fecho convexo* H de um conjunto arbitrário S é o menor conjunto convexo que contiver S . A diferença entre os conjuntos $H - S$ é chamada *deficiência convexa* D do conjunto S . Para imaginar como estes conceitos podem ser utilizados, considere a

Figura 2, que mostra um objeto (conjunto S) e sua deficiência convexa (regiões sombreadas): GONZALES e WOODS (2000, p.350).

FIGURA 11 – REGIÃO (S) E SUA DEFICIÊNCIA CONVEXA REPRESENTADA COMO ÁREAS NA COR CINZA.



FONTE: ADAPTADO DE GONZALES E WOODS, 2000.

2.4 REDES NEURAIS

A integração da informação espectral dos segmentos e sua forma pode ser realizada de diferentes maneiras como, por exemplo, usando abordagens estatísticas, lógica *Fuzzy* ou redes neurais. Atualmente o amadurecimento das técnicas computacionais inteligentes, como Sistemas Especialistas, Lógica *Fuzzy* e Redes Neurais, têm permitido novas abordagens para problemas de processamento de imagens onde processos tradicionais não mais são eficientes.

De acordo com HAYKIN (2001, p.27), o trabalho em Redes Neurais Artificiais (RNA's), usualmente denominadas "Redes Neurais", tem sido motivado pelo reconhecimento de que o cérebro humano processa informações de uma forma inteiramente diferente do computador digital convencional. O cérebro é um sistema de processamento de informação altamente complexo, não-linear e paralelo, que possui a capacidade de organizar seus constituintes estruturais, conhecidos por neurônios, de forma a realizar certos processamentos como por exemplo, reconhecimento de padrões, percepção e controle motor, muito mais rapidamente do que o mais ágil computador hoje existente.

Conforme NISHIDA (1998), entender o funcionamento do cérebro humano sempre foi um desafio à comunidade científica. Sua agilidade e eficiência constituem o objetivo a ser alcançado por sistemas da Inteligência Artificial (IA), que baseia seus estudos na simulação computacional de aspectos da inteligência humana.

As RNA's se assemelham ao cérebro em dois aspectos:

- 1) o conhecimento é adquirido pela rede a partir de seu ambiente através de um processo de aprendizado;
- 2) forças de conexão entre neurônios, conhecidas como pesos sinápticos, são utilizadas para armazenar o conhecimento adquirido (HAYKIN, 2001, p.28).

2.4.1 O SISTEMA NERVOSO

O sistema nervoso, juntamente com o sistema endócrino, produz a maioria das funções de controle do corpo. Ele é formado, basicamente, por células nervosas, que se interconectam de forma específica e precisa, formando os chamados circuitos neurais. Através desses circuitos, o organismo é capaz de produzir respostas estereotipadas que constituem os comportamentos fixos e invariantes (por exemplo, os reflexos), ou então, produzir comportamentos variáveis em maior ou menor grau (LOESCH e SARI, 1996, p.8-11).

Conforme HAYKIN (2001, p.29), todo ser vivo dotado de um sistema nervoso é capaz de modificar o seu comportamento em função de experiências passadas. Essa modificação comportamental é chamada de aprendizado, e ocorre no sistema nervoso através da propriedade chamada plasticidade cerebral. Esta propriedade permite que o sistema nervoso em desenvolvimento se adapte ao seu meio ambiente. Em um cérebro adulto, a plasticidade pode ser atribuída a dois mecanismos: a criação de novas conexões sinápticas entre os neurônios e a modificação das sinapses existentes.

NELSON & ILLINGWORTH (1991), definem as RNA's como sendo técnicas inovadoras de processamento de informações, que têm como objetivo básico, simular, em computadores, o funcionamento do sistema nervoso biológico.

2.4.1.1 O Neurônio Biológico

Como as RNA's são baseadas na estrutura do cérebro (neurônio), serão mostradas as principais características de um neurônio biológico, para uma posterior compreensão, a respeito do funcionamento das RNA's.

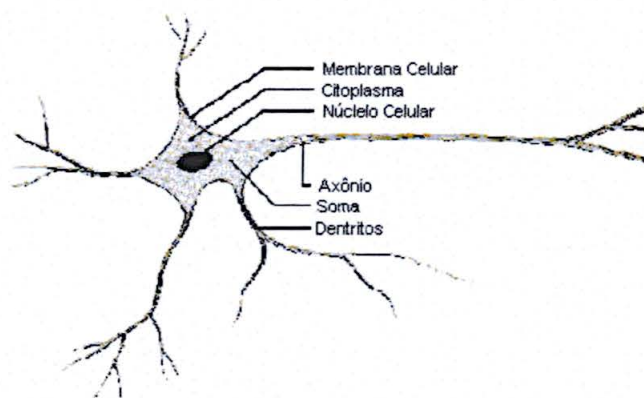
O neurônio biológico é uma célula que pode ser dividida em três seções: o corpo da célula, os dendritos e o axônio, cada uma com funções específicas, porém complementares.

De acordo com LOESCH e SARI (1996, p.8), o corpo do neurônio mede apenas alguns milésimos de milímetros e os dendritos apresentam poucos milímetros de comprimento. O axônio pode ser mais longo. Os dendritos têm como função receber informações, ou impulsos nervosos, oriundas de outros neurônios, e conduzi-las até o corpo celular. Ali, a informação é processada e novos impulsos são gerados. Estes impulsos são transmitidos a outros neurônios (nodos), passando pelo axônio e atingindo os dendritos dos neurônios seguintes.

O ponto de contato entre a terminação axônica de um neurônio e o dendrito de outro é chamado sinapse. É pelas sinapses que os nodos se unem funcionalmente, formando as redes neurais. As sinapses funcionam como válvulas, sendo capazes de controlar a transmissão de impulsos, isto é, o fluxo da informação entre os nodos na rede neural. O efeito das sinapses é variável, é esta variação que dá ao neurônio capacidade de adaptação (FABRI, 1999, p.4).

Na Figura 12 está ilustrada, de forma simplificada, os componentes do neurônio.

FIGURA 12 - COMPONENTES DE UM NEURÔNIO BIOLÓGICO



FONTE: FABRI, 1999, p.4.

Os sinais oriundos dos nodos pré-sinápticos são passados para o corpo do neurônio, onde são combinados com outros sinais recebidos pelo neurônio. Se o percentual em um intervalo curto de tempo é suficientemente alto, a célula “dispara”, produzindo um impulso que é transmitido para as células seguintes (nodos pós-sinápticos). Este sistema simples é responsável pela maioria das funções realizadas pelo nosso cérebro. A capacidade de solucionar funções complexas surge com a operação em paralelo de todos os 1011 nodos do nosso cérebro (FABRI, 1999, p.4).

2.4.1.2 Redes Neurais Artificiais – O Modelo MCP (McCulloch e Pitts)

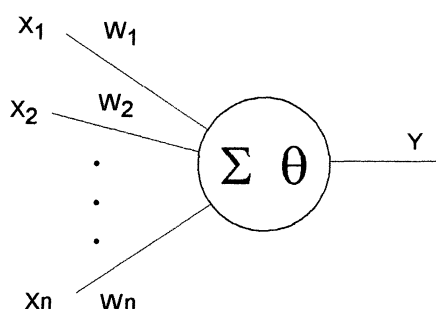
As redes neurais artificiais foram desenvolvidas, originalmente, na década de 40, pelo neurofisiologista Warren McCulloch e pelo matemático Walter Pitts, os quais dentro do espírito cibernético, fizeram uma analogia entre células nervosas vivas e o processo eletrônico num trabalho publicado sobre "neurônios formais". O trabalho consistia num modelo de resistores variáveis e amplificadores representando conexões sinápticas de um neurônio biológico (AZEVEDO *et al.*, 2000, p.4).

Desde então, mais enfaticamente a partir da década 80, diversos modelos de redes neurais artificiais têm surgido com o propósito de aperfeiçoar e aplicar esta tecnologia. Algumas destas propostas tendem a aperfeiçoar mecanismos internos da rede neural para aplicação na indústria e negócios, outras procuram aproximá-las ainda

mais dos modelos biológicos originais.

Conforme DYMINSKI (2000), o modelo de neurônio proposto por McCulloch e Pitts (MCP) em 1943 é uma simplificação do que se sabia a respeito do neurônio biológico naquela época. A sua descrição matemática resultou em um modelo com n terminais de entrada x_1, x_2, \dots, x_n (que representam os dendritos), e apenas um terminal de saída y (que representa o axônio). Uma descrição do modelo está ilustrado na Figura 13.

FIGURA 13 - NEURÔNIO DE MCCULLOCH E PITTS.



FONTE: DYMINSKI, 2000.

Para emular o comportamento das sinapses, os terminais de entrada do neurônio têm pesos acoplados (w_1, w_2, \dots, w_n), cujos valores podem ser positivos ou negativos, dependendo das sinapses correspondentes serem inibitórias ou excitatórias. O efeito de uma sinapse particular " i " no neurônio pós-sináptico é dado por: " $x_i w_i$ ". Os pesos determinam "em que grau" o neurônio deve considerar sinais de disparo que ocorrem naquela conexão (FABRI, 1999, p.6).

De acordo com DYMINSKI (2000), um neurônio dispara quando a soma dos impulsos que ele recebe ultrapassa o seu limiar de excitação chamado de *threshold*. O corpo do neurônio, por sua vez, é emulado por um mecanismo simples que faz a soma dos valores " $x_i w_i$ " recebidos pelo neurônio (soma ponderada) e decide se o neurônio deve ou não disparar (saída igual a 1 ou a 0) comparando a soma obtida ao limiar ou *threshold* do neurônio. No modelo MCP, a ativação do neurônio é obtida através da aplicação de uma "função de ativação", que ativa a saída ou não, dependendo do valor

da soma ponderada das suas entradas. Na descrição original do modelo MCP, a função de ativação é dada pela função limiar descrita na equação 26, a seguir. O nodo MCP terá então sua saída ativa quando:

$$\sum_{i=1}^n x_i w_i \geq \theta \quad (23)$$

onde " n " é o número de entradas do neurônio, " w_i " é o peso associado à entrada " x_i ", e " θ " é o limiar (*threshold*) do neurônio.

McCulloch e Pitts simplificaram seu modelo assumindo que os nodos em cada camada da rede disparam sincronicamente, isto é, que todos os nodos são avaliados ao mesmo tempo e também que as entradas em um instante de tempo " t " produzem a sua saída no tempo " $t + 1$ ". Em sistemas biológicos, sabe-se que não existe um mecanismo para sincronizar as ações de nodos, nem há restrição para que as suas saídas sejam ativadas em tempos discretos como no modelo MCP. Sabe-se também que o valor da próxima saída dos nodos biológicos depende enormemente das ativações dos neurônios anteriores, já que até mesmo os neuro-transmissores liberados anteriormente levam algum tempo para se recombinarem, influenciando, assim, as ativações seguintes. Finalmente, pode-se levantar algumas limitações na descrição do modelo MCP original, como:

- 1) redes MCP com apenas uma camada conseguem implementar somente classes linearmente separáveis;
- 2) pesos negativos são mais adequados para representar disparos inibidores;
- 3) o modelo foi proposto com pesos fixos, não ajustáveis. Não há a presença de um algoritmo de aprendizado DYMINSKI (2000).

2.4.1.2.1 Funções de Ativação

A partir do modelo proposto por McCulloch e Pitts foram derivados vários outros modelos que permitem a produção de uma saída qualquer, não necessariamente zero ou um, e com diferentes funções de ativação. A Figura 14 ilustra, graficamente, quatro funções de ativação diferentes: a função linear, a função rampa, a função degrau (*step*) e a função sigmoidal (HAYKIN, 2001, p.39 e TAFNER, XERES e FILHO, 1996, p.59).

Conforme FABRI (1999, p.8), a função de ativação linear, mostrada na Figura 14a, é definida pela equação abaixo:

$$y = \alpha x \quad (24)$$

onde α é um número real que define a saída linear para os valores de entrada, y é a saída e x é a entrada.

A função linear pode ser restringida para produzir valores constantes em uma faixa $[-\gamma, +\gamma]$, e neste caso, passa a ser a função rampa mostrada, graficamente, na Figura 14b, definida pela notação abaixo:

$$y = \begin{cases} +\gamma & \text{se } x \geq \gamma \\ x & \text{se } x = \gamma \\ -\gamma & \text{se } x \leq -\gamma \end{cases} \quad (25)$$

Os valores máximo e mínimo da saída são $-\gamma, +\gamma$. A função rampa é usada geralmente, como uma função linear simplificada.

A função "degrau", ilustrada na Figura 14c, representa uma resposta binária, pois a função produz a saída $+\gamma$ para valores de x maiores que zero, caso contrário a função produz o valor de $-\gamma$. A função degrau é definida por:

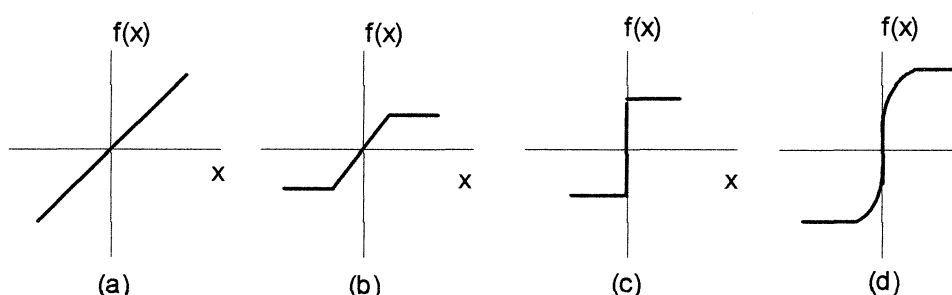
$$y = \begin{cases} +\gamma & \text{se } x > 0 \\ -\gamma & \text{se } x \leq 0 \end{cases} \quad (26)$$

A função sigmoidal, conhecida também como curvas "S", ilustrada na Figura 14d, é uma função semi-linear. É possível definir várias funções sigmoidais. Uma das funções sigmoidais mais importantes é a função logística definida por:

$$y = \frac{1}{1 + e^{-x/T}} \quad (27)$$

onde parâmetro T determina a suavidade de curva.

FIGURA 14 - ALGUMAS FUNÇÕES DE ATIVAÇÃO.



FONTE: FABRI, 1999.

2.4.1.3 Arquitetura das Redes Neurais

Conforme TAFNER, XERES e FILHO (1996, p.62), uma rede neural é constituída por diversos neurônios. Para se formar uma rede neural, diversos neurônios são conectados entre si, sendo que a maneira como os mesmos são dispostos caracteriza a arquitetura da rede.

Arquiteturas neurais são tipicamente organizadas em camadas, com unidades que podem estar conectadas às unidades da camada posterior.

Usualmente as camadas são classificadas em três grupos (TAFNER, XERES e FILHO, 1996, p.62):

- camada de Entrada: onde os padrões são apresentados à rede;
- camadas Intermediárias ou Escondidas: onde é feita a maior parte do processamento, através das conexões ponderadas; podem ser consideradas como extratoras de características;

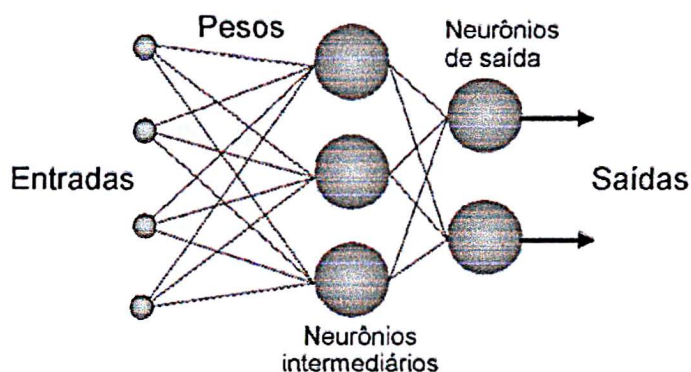
- camada de Saída: onde o resultado final é concluído e apresentado.

O tipo de rede chamada de “multicamadas *feedforward*” caracteriza-se por apresentar neurônios, total ou parcialmente conectados por sinapses (pesos), organizados em camadas. As entradas alimentam a rede e são processadas até alcançarem a camada de saída, sem realimentação. Por isso são caracterizadas como *feedforward* (alimentação e propagação para frente). É o tipo de rede mais utilizado em aplicações de diversas áreas da engenharia (DYMINSKI, 2000).

Na Figura 15, pode ser observada uma rede neural totalmente conectada, com 4 entradas e 2 saídas e que possui uma camada escondida interna, localizada entre o nó e a camada de saída: esta possuindo 3 neurônios.

Sabe-se que a maior parte dos problemas de engenharia podem ser resolvidos com redes de uma ou duas camadas escondidas apenas (HAYKIN, 2001, p.46).

FIGURA 15 - EXEMPLO DE UMA REDE NEURAL.



Conforme HAYKIN (2001, p.46-48), a estrutura básica de uma RNA pode apresentar-se da seguinte forma:

- a) uma camada inicial de unidades sensoras aos estímulos de entrada, que não é composta por neurônios;
- b) uma ou mais camadas intermediárias de neurônios interligados (formando a estrutura central de processamento da rede neural);
- c) e uma camada de neurônios de saída, conectada aos neurônios da(s) camada(s) escondida(s), gerando a resposta final da RNA.

A Figura 15 ilustra a estrutura básica de uma rede neural artificial. As entradas

da rede representam a informação existente. Elas alimentam a primeira camada de neurônios que através do cálculo executado pela aplicação da função de transferência, geram repostas que alimentam a camada seguinte, e assim sucessivamente até atingirem a camada de saída.

O número de neurônios e camadas em uma dada estrutura é função do problema a ser resolvido. A camada de saída, em um problema de classificação de dados, pode representar o número de classes para as quais os padrões podem ser designados. Na classificação de imagens de satélite, por exemplo, se os padrões de entrada estão localizados em uma área onde há a existência de três classes, "Vegetação", "Área Urbana" e "Água", três neurônios de saída serão necessários (NISHIDA, 1998).

2.4.1.4 Conjunto de Dados

Analogamente a outros métodos estatísticos, a escolha do conjunto de dados que servirá para alimentar o modelo neural é muito importante. Como em todo o procedimento de classificação, este conjunto de dados deverá cobrir todo o domínio de interesse do problema a ser solucionado (DYMINSKI, 2000).

Na construção de um modelo neural, deve-se inicialmente analisar o problema a ser modelado, quais suas características mais importante, os fatos relevantes que podem influenciá-lo como também, as respostas que se querem obter como saídas da rede neural.

2.4.1.5 Aprendizado de uma RNA

As RNA's possuem a capacidade de aprender através exemplos e fazer interpolações e extrapolações do que aprenderam. No aprendizado conexionista não se procura obter regras como na abordagem simbólica da Inteligência Artificial (IA), mas determinar a intensidade das conexões entre os neurônios. O conjunto de procedimentos para adaptar os parâmetros de uma RNA, é chamado de algoritmo de aprendizado. Como era de se esperar, não há um único algoritmo de aprendizado. Existem algoritmos que, basicamente, diferem pela maneira através da qual o ajuste dos pesos é feito (DYMINSKI, 2000).

Na fase de aprendizagem, a rede extrai características relevantes de padrões de informações apresentados para a mesma, criando assim uma representação própria para o problema. De acordo com TAFNER, XERES e FILHO (1996, p.65), diversos métodos para treinamento de redes foram desenvolvidos, podendo estes serem agrupados em dois métodos principais: aprendizado supervisionado (com professor) e aprendizado não supervisionado (sem professor).

2.4.1.5.1 Aprendizado Supervisionado

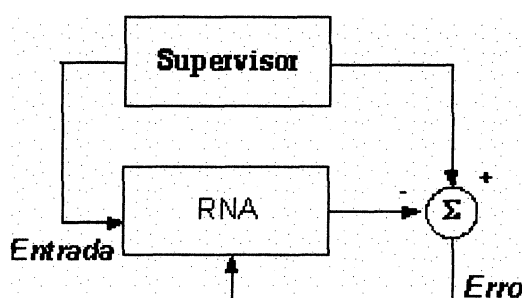
Conforme AZEVEDO, BRASIL e OLIVEIRA (2000, p.20), este método de aprendizado é o mais comum no treinamento das RNA's, e a mesma é treinada com o auxílio de um professor, ou de um supervisor. Para tanto, a rede deverá possuir pares de entrada e saída, ou seja, um conjunto de entradas e um conjunto com as saídas desejadas para cada entrada. O objetivo é ajustar os parâmetros da rede, de forma a encontrar uma ligação entre os pares de entrada e saída fornecidos. A Figura 16 ilustra o mecanismo de *aprendizado supervisionado*.

Toda entrada apresentada à rede, deverá ser verificada se a saída obtida (gerada a partir dos cálculos efetuados com os pesos que a rede possui) confere com a saída

desejada àquela entrada. Sendo diferente, a rede deverá ajustar os pesos das conexões para minimizar o erro. A minimização da diferença é incremental, já que pequenos ajustes são feitos nos pesos a cada etapa de treinamento, de tal forma que estes caminhem, se possível, para uma solução (TAFNER, XERES e FILHO, 1996, p.65-72).

O ajuste dos pesos deverá ser repetido iterativamente com todo o conjunto de treinamento (entradas e saídas), até que a taxa de acerto esteja dentro de uma faixa considerada satisfatória.

FIGURA 16 - APRENDIZADO SUPERVISIONADO.



FONTE: FABRI, 1999.

O exemplo mais conhecido de algoritmo de aprendizado supervisionado para redes de múltiplas camadas, é o algoritmo *Backpropagation* (AZEVEDO, BRASIL e OLIVEIRA, 2000, p.16).

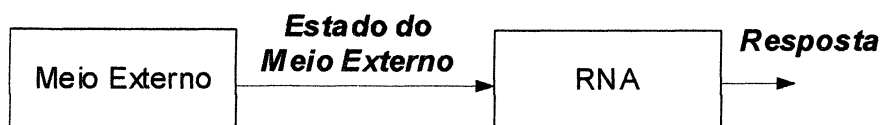
2.4.1.5.2 Aprendizado Não-Supervisionado

No Aprendizado Não Supervisionado, como o próprio nome sugere, não há um “professor” ou supervisor para acompanhar o processo de aprendizado. Este método é ilustrado na Figura 17. Para este tipo de aprendizado, somente os padrões de entrada estão disponíveis para rede, não há uma saída desejada (AZEVEDO, BRASIL e OLIVEIRA, 2000, p.21).

Conforme TAFNER, XERES e FILHO (1996, p.65), esse tipo de rede utiliza os neurônios como classificadores, e os dados de entrada como os elementos para classificação. O processo de classificação fica a encargo da rede neural e o seu

algoritmo de aprendizado.

FIGURA 17 - APRENDIZADO NÃO-SUPERVISIONADO.



FONTE: FABRI, 1999.

2.4.1.6 Treinamento

Após a escolha do tipo e arquitetura da rede a ser utilizada na construção do modelo, definidos os parâmetros que servirão de entradas à rede e as saídas que se quer obter e coletados os dados referentes ao fenômeno a ser modelado, pode-se iniciar o processo de treinamento da rede neural.

De acordo com DYMINSKI (2000), deve-se dividir o conjunto de dados em conjunto de treinamento e de testes. O conjunto de treinamento (também chamado de *In Sample*) será composto de exemplos que serão utilizados na fase de aprendizado da rede neural. Já o conjunto de teste (*Out of Sample*) será composto de exemplos que serão utilizados na tarefa de se verificar a capacidade de generalização da rede. Cabe salientar que os exemplos de teste não deverão ser utilizados no treinamento da rede, ou seja, o conjunto de treinamento e o conjunto de teste são formados por exemplos disjuntos.

O treinamento é um processo iterativo, onde os fatos (conjunto de entrada) são apresentados à rede continuamente até que, então, acerte todos os fatos, ou pelo menos uma razoável percentagem deles (TAFNER, XERES e FILHO, 1996, p.77).

O objetivo da fase de treinamento é ajustar da melhor maneira possível os valores dos pesos sinápticos da rede, procurando fazer com que as entradas fornecidas pela rede estejam bastante próximas dos valores reais de saída correspondentes a cada um dos exemplos apresentados, sem que se perca sua capacidade de *generalização*.

De acordo com HAYKIN (2001, p.28), “a *generalização* refere-se ao fato de

produzir saídas adequadas para entradas que não estavam presentes durante o treinamento (*aprendizagem*”).

Existem diversos algoritmos para se treinar as redes neurais (HAYKIN, 2001), sendo o mais comumente utilizado e adequado para reconhecimento de padrões e classificação o de *Backpropagation* (retro-propagação do erro) (LOESCH e SARI, 1996, p.46).

2.4.1.6.1 O Algoritmo Backpropagation

O treinamento de uma RNA através do algoritmo *Backpropagation* consiste basicamente no seguinte processo:

- a) as entradas dos exemplos do conjunto de treinamento são apresentadas à rede;
- b) estes dados são processados pela rede: em cada camada, através da multiplicação pelos pesos sinápticos, posterior somatório e aplicação da função de ativação, são fornecidas as saídas relacionadas a entradas apresentadas;
- c) as saídas fornecidas pela rede são comparadas com os valores reais obtidos dos experimentos e, desta variação, é calculado o erro;
- d) com o valor do erro, o ajuste para os pesos da última camada é calculado. O erro é então “retro-propagado” na rede, corrigindo os pesos sinápticos das camadas escondidas, visando um melhor ajuste da rede neural ao fenômeno a ser modelado.

Este processo iterativo deverá acontecer até que o erro seja aceitável, ou seja, que a rede tenha aprendido a tarefa a ser realizada.

2.4.1.7 Exemplos de Aplicação das Redes Neurais

São várias as aplicações das redes neurais nesta área: PERELMUTER *et al.* (1995) utilizaram as Redes Neurais Artificiais no reconhecimento de caracteres como peças mecânicas, chaves e canetas presentes em imagens digitais, onde o objetivo era

o de avaliar o desempenho das redes neurais quando comparadas aos modelos tradicionais de classificação, e uma das vantagens apresentadas foi a de que o classificador neural não há a necessidade de uma separação preliminar das peças existentes nas imagens. As peças foram apresentadas à rede sem nenhuma pré-classificação.

SANTOS *et al.*(1999), utilizaram uma rede *feedforward* no monitoramento da condição de ferramentas de corte. Eles utilizaram as redes neurais para reconhecer a ocorrência do desgaste da ferramenta numa operação de fresado, através das informações geradas pelos sensores instalados num centro de usinagem foi acoplado ao sistema de monitoramento, um sistema de tomada de decisões, o qual apropriadamente interpretará as informações de entrada do sensor associando-as aos padrões do estado da ferramenta. Para a identificação de ferramentas novas e desgastadas, os resultados obtidos foram de 95% a 100%.

LIBERMAN (1997) empregou as redes neurais com parâmetros de textura na classificação de imagens digitais. NISHIDA (1998) utilizou uma rede neural artificial com Função de Base Radial para classificação de imagens Landsat 5 da região de Florianópolis.

CAPÍTULO 3

MATERIAIS E MÉTODOS

3.1 MATERIAIS EMPREGADOS

3.1.1 IMAGEM UTILIZADA

A imagem utilizada neste trabalho foi um recorte de uma cena do sensor *IKONOS* da região de Matinhos, município pertencente ao Litoral Paranaense. Esta imagem multiespectral já registrada possui uma resolução espacial de 1 m, resultado de uma fusão das imagens multiespectrais (resolução espacial de 4 m) e pancromática (resolução espacial de 1 m).

As principais características observadas da área da imagem utilizada são: área de asfalto, sendo estas localizadas junto à nova Rodoviária; ruas e solo exposto compostos de areia; vegetação de alta e baixa densidade; construções com uma grande variação de tipos de materiais, e; sombra resultante das construções e árvores.

A Figura 18 mostra um recorte da imagem *IKONOS*. A imagem captada pelo sensor *IKONOS* permite identificar maiores detalhes da paisagem, mas este aumento na resolução espacial também ocasiona maior confusão em termos espectrais. Objetos diferentes podem ter a mesma resposta espectral, enquanto pode também ocorrer que objetos da mesma classe apareçam com cores diferentes. Por exemplo, existe uma grande variabilidade dos tipos de materiais que constituem as construções, o que causa uma certa dificuldade em se distinguir áreas cobertas por areia e construções, pois alguns telhados possuem aparência espectral similar à areia. Isto também pode ser percebido ao verificar-se a confusão entre ruas, asfalto e telhados de mesma cor.

FIGURA 18 – IMAGEM IKONOS DE MATINHOS.



3.2 METODOLOGIA

Considerando a grande confusão espectral verificada nas imagens de alta resolução, torna-se evidente a necessidade de incluir novas informações no processo de classificação, a fim de aumentar a separabilidade entre as classes de objetos. Neste trabalho, utiliza-se a forma dos segmentos como informação auxiliar.

O reconhecimento de objetos em imagens de alta resolução espacial demanda uma interpretação mais apurada, ou seja, não mais a interpretação dos *pixels* mas sim a de um conjunto deles, levando em consideração todo o seu contexto. Esta interpretação ou classificação é realizada então através da análise *orientada ao objeto*.

A seguir são apresentados os procedimentos implantados e testados neste trabalho para a este tipo de classificação.

3.2.1 SEGMENTAÇÃO DA IMAGEM

A classificação baseada na metodologia de dados orientados *ao objeto* baseia-se na análise de agregações de pixels adjacentes similares. O princípio desta análise

considera não apenas as características espectrais de um único *pixel*, mas as características espectrais de um grupo de *pixels*, bem como suas características espaciais e contextuais. Para isto, torna-se necessário dividir a imagem em segmentos, considerados objetos na imagem, os quais são extraídos da imagem original por meio de técnicas de segmentação. Tais objetos são posteriormente analisados como unidades, sendo possível a análise de fatores de forma, textura e outras variáveis espaciais que não podem ser consideradas em uma análise *pixel a pixel*.

Para a segmentação da imagem utilizada neste trabalho, optou-se pelo algoritmo de crescimento de regiões do *Software eCognition*, disponível no mercado desde outubro de 2000. Este algoritmo divide inicialmente a imagem em segmentos elementares e posteriormente combina estes segmentos para formar regiões maiores com características espectrais similares. Com isto, é possível estimar os limites de objetos homogêneos da imagem, especificando ainda o tamanho esperado dos objetos (ECOGNITION, 2001).

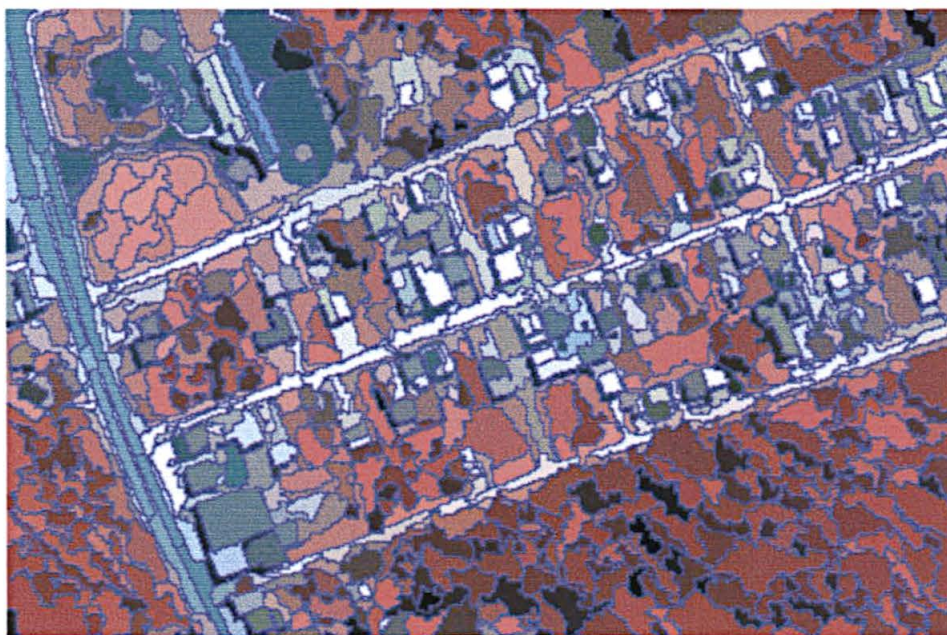
Para a segmentação neste *software*, há a necessidade da definição de vários parâmetros, entre os quais estão os principais: *image layers*, *scale parameter* e *homogeneity criterion* (ECOGNITION, 2001).

1. *image layers*: este campo refere-se à escolha das bandas espectrais a serem utilizadas durante o processo de segmentação;
2. *scale parameter*: o parâmetro de escala determina a máxima heterogeneidade permitida para os objetos, e o mesmo influencia o tamanho médio dos objetos gerados. Este parâmetro reflete dois critérios de uniformidade: Uniformidade espectral, representada pela variação dos valores digitais na região, e a uniformidade espacial, representada através de um índice que descreve a compacidade da região e a sinuosidade de seus contornos (ECOGNITION, 2001).
3. *homogeneity criterion*: o critério de homogeneidade é utilizado para especificar o peso que deve ser atribuído à uniformidade espectral e à

uniformidade espacial no cômputo do parâmetro de escala. Quanto maior for o peso atribuído ao parâmetro de forma, menor é a influência da homogeneidade espectral na geração dos objetos. Quando o critério de forma é superior a zero, o usuário pode determinar se os objetos deverão ser mais compactos ou mais suavizados.

Realizaram-se vários testes com a variação destes parâmetros, principalmente em relação ao parâmetro de escala, a fim de se obter um tamanho ideal para os objetos. Com isto pôde-se obter uma imagem segmentada (Figura 19), a qual foi base para a posterior classificação *orientada a regiões*.

FIGURA 19 – IMAGEM RESULTANTE DA SEGMENTAÇÃO NO SOFTWARE ECOGNITION.



3.2.2 CODIFICAÇÃO DE CADA SEGMENTO DA IMAGEM

O resultado da segmentação no *Ecognition* foi então exportado para ser utilizado no ambiente MATLAB. Como a imagem “exportada” no formato *TIFF* perde a associação com a tabela descritiva dos segmentos, houve a necessidade de que todos

os segmentos da imagem fossem novamente “rotulados”, ou seja, a cada objeto ou região da imagem gerada foi atribuído um número identificador único, ou um código.

Os dados de entrada foram organizados sob forma de uma tabela, onde cada coluna da mesma possui: número do segmento, valor espectral na banda do infravermelho próximo (*IVP*), na banda do vermelho (*R*), e na banda do verde (*G*), valor do parâmetro de forma. Esta tabela seria posteriormente utilizada como entrada na classificação por redes neurais dentro do ambiente *MATLAB*. A mesma pode ser observada no capítulo 4, “Resultados e Discussões”.

Um programa na linguagem *MATLAB* foi elaborado para a codificação dos objetos da imagem segmentada pelo eCognition, visto que na exportação, os identificadores do eCognition são perdidos. Este programa é baseado no princípio do algoritmo de crescimento de regiões que, neste caso, são completamente uniformes. No resultado, os pixels de um mesmo segmento foram representados usando o mesmo identificador, aquele correspondente à região.

Para a geração da codificação dos objetos da imagem, foi realizado um programa na linguagem *MATLAB*, onde para tanto se optou por uma varredura da imagem, *pixel a pixel*, onde o agrupamento dos *pixels* considerados de mesmo código foi realizado através de critérios de homogeneidade espectral entre os mesmos. Então se gerou uma imagem (Figura 20) codificada em formato *matricial*, onde é possível conhecer as coordenadas *linha e coluna* de segmento da imagem.

FIGURA 20 – IMAGENS DE SEGMENTOS CODIFICADOS.

1	1	1	1	1	1
1	1	1	2	2	2
3	3	3	4	4	5
3	3	3	4	5	5
3	3	4	4	5	5
6	6	6	4	5	5
6	6	6	6	6	5
6	6	6	6	7	7
8	8	8	7	7	7
8	8	7	7	7	7
9	9	7	7	7	7
9	9	9	10	10	10
9	9	9	9	10	10

Uma vez identificadas às regiões na imagem, o seguinte passo foi calcular os valores médios da resposta espectral de cada uma delas. Para isto, os valores digitais dos pixels de uma região foram lidos na imagem original. Assim, uma tabela com 4 colunas foi elaborada, contendo o identificador e os três valores médios de cada região.

3.2.3 DEFINIÇÃO DOS PARÂMETROS NECESSÁRIOS À CLASSIFICAÇÃO POR REDES NEURAIS ARTIFICIAIS

3.2.3.1 Parâmetros da Arquitetura da Rede

A configuração da rede é um processo difícil e importante, e a acurácia e tempo de processamento é dependente da alteração destes parâmetros (JARVIS e STUART, 1996).

Os parâmetros básicos necessários para a definição da arquitetura de uma rede neural são: tipo de rede e número de neurônios das camadas de entrada, intermediária ou escondida da camada de saída.

Como o objetivo do trabalho é a classificação de padrões, o tipo de rede utilizada foi a *Feedforward*, que possui o algoritmo denominado *Backpropagation*, já que a mesma é a recomendada para a classificação supervisionada de imagens de sensoriamento remoto (LEES, 1996; JARVIS e STUART, 1996; MURNION, 1996; LOESCH e SARI, 1996).

O número de neurônios de entrada foi bastante variável durante o trabalho, já que o mesmo é função dos dados de entrada, e como vários parâmetros de forma foram analisados, o número de entradas não foi constante ao longo dos testes realizados. Os vários parâmetros foram analisados 2 a 2, 3 a 3 e assim por diante, de acordo com a sua contribuição no resultado da classificação, e os menos relevantes ao longo das

tentativas foram sendo descartados nas combinações.

As entradas foram apresentadas variando de no mínimo 3, onde apenas os parâmetros espectrais das bandas (*IVP-R-G*) foram apresentados à rede; e o maior número de entradas foi 12, onde além dos 3 parâmetros espectrais mais 9 parâmetros de forma foram utilizados como entrada para a rede neural.

O número de neurônios da camada de saída da rede é função do número de classes da imagem resultante da classificação, e a mesma é fornecida para rede pelo usuário. No caso deste trabalho o número de neurônios da camada de saída foi 8, já que foram identificadas 8 classes na imagem utilizada.

Para a camada escondida, é recomendável fazer alguns testes variando o número de neurônios. Se a camada escondida possui poucos neurônios, a rede pode ficar incapaz de criar limites de decisão muito complexos, enquanto que um número excessivo dos mesmos acarreta perda da capacidade de generalização da rede (LOESCH e SARI, 1996).

Muitos autores estudaram o efeito que do número de neurônios na camada escondida causa na acurácia e tempo de convergência, e confirmam a dificuldade neste processo de configuração deste parâmetro (JARVIS e STUART, 1996; GERMAN e GAHEGAN, 1996; FITZGERALD e LEES, 1996; MURNION, 1996), ou então encontraram *regras* de melhor arquitetura para a classificação de imagens *LANDSAT*. Por exemplo, algumas recomendações são que o número de neurônios na camada escondida deve ser maior que o número de neurônios de entrada, ou que o número de camadas escondidas deve ser 3 vezes o número de amostras de treinamento. Até mesmo expressões para se encontrar o tamanho ideal de amostras de treinamento para a arquitetura definida foram propostas (JARVIS e STUART, 1996; GERMAN e GAHEGAN, 1996; FITZGERALD e LEES, 1996; MURNION, 1996; PAOLA e SCHOWENGERDT, 1997).

GERMAN e GAHEGAN (1996); FITZGERALD e LEES (1996), apresentaram uma arquitetura ideal para a classificação temporal de imagens *LANDSAT*. Entretanto

mesmo tratando-se do mesmo tipo de imagens, ou seja, imagens coletadas pelo mesmo sensor, as “receitas” observadas apresentaram-se distintas.

Por este motivo optou-se por utilizar duas imagens “testes” para o auxílio à obtenção da melhor arquitetura para a classificação por redes neurais, já que se faz necessário o encontro do número de camadas escondidas e o número de neurônios em cada uma destas camadas, ou seja, criar uma *regra padrão* para imagens *IKONOS*.

Uma das imagens é uma imagem sintética (Figura 21), criada com o intuito de simular uma imagem com características controladas. Também foi utilizada uma imagem mais simples do sensor *IKONOS* (Figura 22, obtida em: <http://www.engesat.com.br/catalogo/ikonos>). Ambas imagens foram utilizadas para se obter uma relação entre o número de amostras utilizadas e número de neurônios utilizados na camada escondida.

Para a imagem sintética, com 28 segmentos, apenas uma amostra foi utilizada para cada classe e obteve-se uma classificação com 100% de acerto para 1 camada escondida com 8 neurônios. Para a imagem mais simples do sensor *IKONOS*, com 140 segmentos, o número de amostras utilizadas foi 4 e o número de neurônios na única camada escondida foi 30.

Esta constatação serviu de base para determinar os parâmetros da rede para classificar a imagem *IKONOS* utilizada na pesquisa, que foram segmentadas em 717 segmentos. Para obter uma primeira estimativa do número de amostras a serem utilizadas, como também o número de neurônios nas camadas escondidas, foi realizado um cálculo simples de *regra de três*. Através do cálculo, os valores encontrados para estes parâmetros foram os obtidos pela *regra de três*, menos 33%. Este valor de percentagem foi encontrado de maneira heurística, já que o valor encontrado através da *regra de três* (de 25 amostras e 206 neurônios para a camada escondida) não foi suficiente, então através de tentativas o valor ideal obtido foi o encontrado pela *regra de três*, menos 33%.

Com estes cálculos, foram encontrados, para uma imagem com 717 segmentos,

os valores de 17 amostras e 138 neurônios para as camadas escondidas. Optou-se por utilizar duas camadas escondidas, já que a imagem apresenta um alto nível de dificuldade de discriminação entre os padrões, e a segunda camada foi composta por 46 neurônios, já que de acordo com alguns autores o número de neurônios da segunda camada deve possuir um terço de neurônios da primeira.

FIGURA 21 – IMAGEM SINTÉTICA.

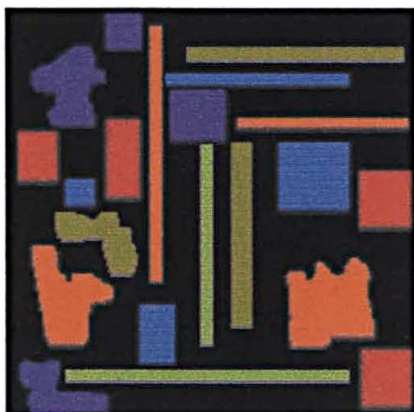


FIGURA 22 – IMAGEM IKONOS SIMPLES.



FONTE: [HTTP://WWW.ENGESAT.COM.BR/CATALOGO/IKONOS](http://www.engesat.com.br/catalogo/ikonos).

3.2.4 CLASSIFICAÇÃO ORIENTADA AO OBJETO UTILIZANDO REDES NEURAIS ARTIFICIAIS

A classificação baseada na metodologia *orientada ao objeto* baseia-se na análise não apenas de um *pixel* isoladamente, mas de regiões previamente segmentadas na imagem. Este tipo de abordagem ganha significado quando a imagem apresenta um grande número de objetos maiores que a resolução espacial. Com a nova geração de sensores de alta resolução, onde as imagens apresentam um alto nível de detalhes, os métodos de classificação *orientados ao objeto* passam a ser objeto de vários estudos. SCHWARZ, STEINMEIER e WASER (2002), utilizaram este método de classificação na detecção de prejuízos causados por tempestades em florestas utilizando dados do satélite *IKONOS* e puderam comprovar que a classificação dos dados de sensores de alta resolução através da metodologia *orientada ao objeto* produz resultados de maior acurácia.

Amostras de treinamento foram coletadas para a realização da classificação. Esta fase define cada uma das classes que se pretende discriminar, levando em consideração sua própria variabilidade na área em estudo. A partir destas amostras, pode-se calcular os valores das variáveis que descrevem/caracterizam a classe. De acordo com CHUVIECO (1990, p.329), pode-se considerar a fase de treinamento das amostras como a coluna vertebral da classificação numérica, pois os resultados da mesma estão diretamente relacionados com a definição prévia das classes. Portanto este conjunto de dados deve cobrir todo o domínio de interesse do problema a ser solucionado.

Foram definidas as principais classes observadas na imagem, sendo estas:

- (1) construções de amianto;
- (2) construções brancas;
- (3) construções de cor verde;
- (4) ruas de asfalto (e Rodoviária);

- (5) areia (ruas e solo exposto);
- (6) vegetação de alta densidade;
- (7) vegetação de baixa densidade, e;
- (8) sombra.

Não foi considerada a classe “não classificada”, já que para a classificação utilizando *objetos*, todos eles foram destinados a uma das classes definidas.

De acordo com LOESCH e SARI (1996, p.50-53), o tamanho do conjunto de treinamento, como regra prática, pode ser calculado tomando-se o dobro do número de neurônios da camada de entrada da rede para se obter o número mínimo de exemplos de treinamento. O mesmo autor recomenda que alguns cuidados devem ser tomados na escolha das amostras: estas devem conter variações de características que a rede deve abstrair, principalmente sem esquecer de incluir casos fronteiros, muito próximos aos limites de decisão, para que a rede possa aprender a estabelecer estes limites.

O número mínimo de amostras necessário para a generalização da rede foi estimado a partir da regra de três proposta acima. Foram coletadas 15 amostras para cada classe. As amostras para uma classificação *orientada a regiões* são distintas daquelas coletadas para uma classificação *pixel a pixel*, já que para esta são coletados grupos de *pixels* onde os mesmos são necessários para um posterior cálculo estatístico. Na coleta de amostras *por regiões*, as amostras são segmentos, ou seja, são coletados vários segmentos como amostra para cada classe, e para cada segmento são calculados a média dos valores espectrais das três bandas, como também os parâmetros de forma e tamanho para o mesmo.

De acordo com HEPNER *et al.* (1990), uma das vantagens das redes neurais é a obtenção de um resultado satisfatório na acurácia da classificação, utilizando-se de um número pequeno de amostras de treinamento. O autor obteve um melhor resultado na classificação por Redes Neurais em comparação à classificação por Máxima Verossimilhança, utilizando um número mínimo de amostras de treinamento.

A rede foi implementada junto ao *software MATLAB*, o qual possui um *toolbox* específico para redes neurais.

A forma de apresentação dos padrões de entrada à rede é muito importante (LOESCH e SARI, 1996, p.50-53).

Conforme TAFNER, XERES e FILHO (1996, p.70), o primeiro ponto a ser discutido na fase de treinamento é que a rede não deve ser treinada 100% para cada fato, mas treinada com o conjunto inteiro, sendo as entradas apresentadas de maneira intercalada. A organização dos dados de entrada segundo as classes é prejudicial, pois a rede não converge, visto que analisando um conjunto consecutivo de pixels de uma mesma classe, ela é especializada para cada classe, perdendo sua adaptação à classe anterior. Assim, quando chegar ao final do conjunto de treinamento, a rede já terá “esquecido” o primeiro fato apresentado.

Quando as amostras são apresentadas de maneira intercalada, pequenos ajustes são feitos em cada iteração, mas em todas as classes. Com isto, a rede não perde sua generalidade. A estratégia de apresentação que foi utilizada é a mais usual (LOESCH e SARI, 1996, p.50-53) e apresenta as amostras de maneira alternada, num esquema de rodízio de classes. No exemplo da tabela 3, a primeira coluna representa a classe e as seguintes os dados de cada segmento, por exemplo, os valores médios das bandas IVP, R, e G. A notação para identificar as classes é a seguinte:

- (0) construções de amianto;
- (1) construções de cor branca;
- (2) construções de cor verde;
- (3) ruas de asfalto (e Rodoviária);
- (4) areia (ruas e solo exposto);
- (5) vegetação de alta densidade;
- (6) vegetação de baixa densidade, e;
- (7) sombra.

TABELA 3 – TABELA DE ENTRADA DAS AMOSTRAS DE TREINAMENTO, APRESENTADA À REDE.

0	1.0995796e+002	8.4986726e+001	8.9823009e+001
1	2.3002174e+002	2.5011957e+002	2.1831522e+002
2	1.4784884e+002	1.7681395e+002	1.3602326e+002
3	4.3653598e+001	1.0762997e+002	1.2206176e+002
4	2.1612313e+002	2.1951119e+002	2.1357090e+002
5	1.2834670e+002	4.4656160e+001	4.8590258e+001
6	2.0752779e+002	9.0897300e+001	1.0263526e+002
7	3.2872340e+000	2.5079787e+001	2.8984043e+001
0	1.5197976e+002	1.4288259e+002	1.4790688e+002
1	1.3844286e+002	1.6742857e+002	1.8595714e+002
2	1.3397581e+002	1.8111290e+002	1.2679032e+002
3	7.0171139e+001	1.2616244e+002	1.4210805e+002
4	1.8007407e+002	1.8744650e+002	1.7751337e+002
5	8.1096774e+001	4.2634409e+001	4.6293907e+001
6	1.5384468e+002	8.5542553e+001	9.2242553e+001
7	5.7217172e+001	2.5601010e+001	2.4843434e+001
...			

Para a montagem deste arquivo, ou tabela, houve a necessidade de se desenvolver vários programas no aplicativo *MATLAB*, e os mesmos são mostrados e descritos a seguir, através da Figura 23.

FIGURA 23 – ESQUEMA DA MONTAGEM DOS PROGRAMAS DE COLETA DE AMOSTRAS.

(1) coleta.dat			(3) tabela.dat		
1 c			segmento	R	G B
3 1			1	4.3653598e+001	1.0762997e+002 1.2206176e+002
64 267			2	2.0255618e+002	2.2808427e+002 2.3908989e+002
124 238			3	1.6308176e+002	1.0816981e+002 1.1677358e+002
144 153			4	1.6476982e+002	8.7063939e+001 9.3056266e+001
41 11			5	1.3879724e+002	1.1320737e+002 1.1991244e+002
117 96			6	2.9985948e+001	8.1817330e+001 9.4967213e+001
6 304			...		
...			7 15	1.4421429e+002	8.4785714e+001 9.0071429e+001
			7 16	1.7110000e+002	1.7940000e+002 1.9000000e+002
			7 17	2.0700000e+002	7.6750000e+001 8.2000000e+001
			(4) amostra.dat		
			classe	R	G B
			0	1.0995796e+002	8.4986726e+001 8.9823009e+001
			1	2.3002174e+002	2.5011957e+002 2.1831522e+002
			2	1.4784884e+002	1.7681395e+002 1.3602326e+002
			3	4.3653598e+001	1.0762997e+002 1.2206176e+002
			4	2.1612313e+002	2.1951119e+002 2.1357090e+002
			5	1.2834670e+002	4.4656160e+001 4.8590258e+001
			6	2.0752779e+002	9.0897300e+001 1.0263526e+002
			7	3.2872340e+000	2.5079787e+001 2.8984043e+001
			0	1.5197976e+002	1.4288259e+002 1.4790688e+002
			1	1.3844286e+002	1.6742857e+002 1.8595714e+002
			2	1.3397581e+002	1.8111290e+002 1.2679032e+002
			3	7.0171139e+001	1.2616244e+002 1.4210805e+002
			4	1.8007407e+002	1.8744650e+002 1.7751337e+002
			5	8.1096774e+001	4.2634409e+001 4.6293907e+001
			6	1.5384468e+002	8.5542553e+001 9.2242553e+001
			7	5.7217172e+001	2.5601010e+001 2.4843434e+001
			...		

A posição das amostras de treinamento (1) é definida em função das coordenadas *linha e coluna* de cada segmento e estas coordenadas são armazenadas num arquivo (*coleta.dat*). A partir destas coordenadas são lidos os valores digitais médios dos segmentos de treinamento da imagem (*matriz.m*) e armazenados numa nova tabela (arquivo *tabela.dat*).

Com isto, a tabela contém o índice da classe e as leituras médias em cada banda. Esta informação pode ainda ser ampliada, aumentando parâmetros espaciais de cada segmento. Para o preenchimento das próximas colunas da tabela de amostras, o programa irá buscar nos arquivos dos parâmetros de forma gerados e buscar, em cada um deles, os correspondentes valores dos segmentos escolhidos. Um exemplo desta tabela com as colunas com parâmetros de forma pode ser visualizada através da Tabela 4.

TABELA 4 – TABELA DE AMOSTRAS DE TREINAMENTO COM ALGUNS PARÂMETROS DE FORMA.

amostra.dat

classe	R	G	B	area	perimetro	compacidade	razao eixos
0	1.0995796e+002	8.4986726e+001	8.9823009e+001	4.5200000e+002	3.3200000e+002	1.9405635e+001	1.5200000e+00
1	2.3002174e+002	2.5011957e+002	2.1831522e+002	9.2000000e+001	5.8000000e+001	2.9097675e+000	3.2100000e+00
2	1.4784884e+002	1.7681395e+002	1.3602326e+002	8.6000000e+001	6.0000000e+001	3.3311500e+000	3.6600000e+00
3	4.3653598e+001	1.0762997e+002	1.2206176e+002	1.8620000e+003	5.5100000e+002	1.2975188e+001	1.0000000e+00
4	2.1612313e+002	2.1951119e+002	2.1357090e+002	8.0400000e+002	3.6600000e+002	1.3258557e+001	2.6000000e+00
5	1.2834670e+002	4.4656160e+001	4.8590258e+001	3.4900000e+002	9.8000000e+001	2.1898626e+000	3.1000000e+00
6	2.0752779e+002	9.0897300e+001	1.0263526e+002	1.8890000e+003	2.8000000e+002	3.3027389e+000	2.0600000e+00
7	3.2872340e+000	2.5079787e+001	2.8984043e+001	1.8800000e+002	9.0000000e+001	3.4286038e+000	5.7000000e+00
0	1.5197976e+002	1.4288259e+002	1.4790688e+002	2.4700000e+002	1.4200000e+002	6.4963568e+000	8.3000000e+00
1	1.3844286e+002	1.6742857e+002	1.8595714e+002	7.0000000e+001	6.0000000e+001	4.0925557e+000	2.7400000e+00
2	1.3397581e+002	1.8111290e+002	1.2679032e+002	1.2400000e+002	7.0000000e+001	3.1445936e+000	1.1300000e+00
3	7.0171139e+001	1.2616244e+002	1.4210805e+002	1.3790000e+003	7.3900000e+002	3.1514814e+001	1.4500000e+00
...							

Na série de testes, o número de parâmetros utilizado foi variado. A variação dos parâmetros de entrada foi feita através do esquema apresentado nas Tabelas 5, 6 e 7, onde para uma posterior análise e comparação, as entradas foram agrupadas quanto ao número de parâmetros utilizados em 10 grupos:

- I – com 3 parâmetros de entrada;
- II – com 4 parâmetros de entrada;
- III – com 5 parâmetros de entrada;
- IV – com 6 parâmetros de entrada;
- V – com 7 parâmetros de entrada;
- VI – com 8 parâmetros de entrada;
- VII – com 9 parâmetros de entrada;
- VIII – com 10 parâmetros de entrada;
- IX – com 11 parâmetros de entrada;
- X - com 12 parâmetros de entrada.

Os 12 parâmetros de entrada, por exemplo, são representados através dos parâmetros espectrais (*IVP*, *R* e *G*) e pelos parâmetros espaciais considerados mais relevantes (*coeficiente de compacidade*, *relação eixo menor e maior da elipse* e os 7 *momentos invariantes*).

TABELA 5 – VARIAÇÃO DO NÚMERO DE PARÂMETROS DE ENTRADA APRESENTADOS À REDE AO LONGO DA PESQUISA.

Nº DE NEURÔNIOS DE ENTRADA GRUPO I	PARÂMETROS UTILIZADOS
3	IVP-R-G
Nº DE NEURÔNIOS DE ENTRADA GRUPO II	PARÂMETROS UTILIZADOS
4	IVP-R-G + ÁREA
4	IVP-R-G + PERÍMETRO
4	IVP-R-G + DEFICIÊNCIA CONVEXA
4	IVP-R-G + COMPACIDADE
4	IVP-R-G + RAZÃO EIXOS
4	IVP-R-G + CÓDIGO DA CADEIA
4	IVP-R-G + 1º MOMENTO
4	IVP-R-G + 2º MOMENTO
4	IVP-R-G + 3º MOMENTO
4	IVP-R-G + 4º MOMENTO
4	IVP-R-G + 5º MOMENTO
4	IVP-R-G + 6º MOMENTO
4	IVP-R-G + 7º MOMENTO
Nº DE NEURÔNIOS DE ENTRADA GRUPO III	PARÂMETROS UTILIZADOS
5	IVP-R-G + ÁREA + COMPACIDADE
5	IVP-R-G + PERÍMETRO + COMPACIDADE
5	IVP-R-G + PERÍMETRO + RAZÃO EIXOS
5	IVP-R-G + ÁREA + RAZÃO EIXOS
5	IVP-R-G + COMPACIDADE + RAZÃO EIXOS
5	IVP-R-G + COMPACIDADE + 1º MOMENTO
Nº DE NEURÔNIOS DE ENTRADA GRUPO IV	PARÂMETROS UTILIZADOS
6	IVP-R-G + ÁREA + PERÍMETRO + COMPACIDADE
6	IVP-R-G + ÁREA + PERÍMETRO + RAZÃO EIXOS
6	IVP-R-G + ÁREA + COMPACIDADE + RAZÃO EIXOS
6	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS
6	IVP-R-G + PERÍMETRO + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA
6	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + DEFIC. CONVEXA
6	IVP-R-G + PERÍMETRO + RAZÃO EIXOS + CÓDIGO DA CADEIA
6	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + CÓDIGO DA CADEIA
6	IVP-R-G + PERÍMETRO + COMPACIDADE + DEF. CONVEXA
6	IVP-R-G + PERÍMETRO + COMPACIDADE + CÓDIGO DA CADEIA
6	IVP-R-G + PERÍMETRO + COMPACIDADE + 1º MOMENTO
6	IVP-R-G + PERÍMETRO + RAZÃO EIXOS + 1º MOMENTO
6	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º MOMENTO
6	IVP-R-G + COMPACIDADE + 1º E 2º MOMENTOS

TABELA 6 - VARIAÇÃO DO NÚMERO DE PARÂMETROS DE ENTRADA APRESENTADOS À REDE AO LONGO DA PESQUISA.

Nº DE NEURÔNIOS DE ENTRADA GRUPO V	PARÂMETROS UTILIZADOS
7	IVP-R-G + ÁREA + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS
7	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA
7	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + CÓDIGO DA CADEIA
7	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + 1º MOMENTO
7	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA + 1º MOMENTO
7	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º E 2º MOMENTOS
7	IVP-R-G + PERÍMETRO + COMPACIDADE + 1º E 2º MOMENTOS
7	IVP-R-G + PERÍMETRO + RAZÃO EIXOS + 1º E 2º MOMENTOS
Nº DE NEURÔNIOS DE ENTRADA GRUPO VI	PARÂMETROS UTILIZADOS
8	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA + CÓDIGO DA CADEIA
8	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + CÓDIGO DA CADEIA + 1º MOMENTO
8	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + 1º E 2º MOMENTOS
8	IVP-R-G + PERÍMETRO + COMPACIDADE + 1º, 2º E 3º MOMENTOS
8	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º E 3º MOMENTOS
Nº DE NEURÔNIOS DE ENTRADA GRUPO VII	PARÂMETROS UTILIZADOS
9	IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA + CÓDIGO DA CADEIA + 1º MOMENTO
9	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º E 3º MOMENTOS
Nº DE NEURÔNIOS DE ENTRADA GRUPO VIII	PARÂMETROS UTILIZADOS
10	IVP-R-G + TODOS OS 7 MOMENTOS
10	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º, 3º, 4º E 5º MOMENTOS

TABELA 7 - VARIAÇÃO DO NÚMERO DE PARÂMETROS DE ENTRADA APRESENTADOS À REDE AO LONGO DA PESQUISA.

Nº DE NEURÔNIOS DE ENTRADA GRUPO IX	PARÂMETROS UTILIZADOS
11	IVP-R-G + COMPACIDADE + TODOS OS 7 MOMENTOS
11	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º, 3º, 4º, 5º E 6º MOMENTOS
Nº DE NEURÔNIOS DE ENTRADA GRUPO X	PARÂMETROS UTILIZADOS
12	IVP-R-G + COMPACIDADE + RAZÃO EIXOS + TODOS OS 7 MOMENTOS
12	IVP-R-G + PERÍMETRO + COMPACIDADE + TODOS OS 7 MOMENTOS

Depois do treinamento das amostras, foi realizada a etapa de simulação da rede neural, ou seja, a classificação de toda a imagem.

3.2.5 AVALIAÇÃO DA CLASSIFICAÇÃO

A precisão pode ser medida tomando-se como base a "verdade de campo", conhecida a partir de trabalhos realizados em campo, da interpretação visual ou ainda de outros levantamentos a respeito da mesma área. Então são comparados os mapas que contenham a verdade terrestre e a os obtidos pela classificação. É necessária a identificação das classes em ambos os mapas, para a formação posterior de uma "matriz de confusão".

Como "verdade de campo", foi realizada uma interpretação visual da imagem. Esta interpretação visual foi efetuada na tela do computador. Polígonos correspondentes a cada classe foram desenhados com ajuda do "mouse" e cada polígono foi identificado como membro de uma das classes propostas. Como era esperado, algumas regiões foram difíceis de serem classificadas, pois sua resposta espectral era confusa. Isto foi verificado, por exemplo, nas regiões de sombra.

Para a geração da "matriz de confusão", ou "matriz de contingência", realizou-se um programa no ambiente *MATLAB*, onde se comparou as amostras de

“verificação” (estas consideradas “verdade de campo”), e a sua correspondência na imagem resultante da classificação. Calculou-se então a *acurácia global* correspondente.

3.2.6 COMPARAÇÃO DO RESULTADO EM RELAÇÃO À CLASSIFICAÇÃO *PIXEL A PIXEL* POR REDES NEURAIS E MÉTODO DA MÁXIMA VEROSSIMILHANÇA

A fim de se avaliar as vantagens da metodologia baseada na segmentação da imagem e a posterior classificação dos segmentos usando redes neurais, optou-se por comparar este resultado com o resultado obtido por métodos tradicionais, como a classificação multiespectral pelo método da máxima verossimilhança e pelas redes neurais pixel a pixel. Ambas as classificações foram realizadas utilizando o *software* de processamento de imagens *ENVI*.

A principal diferença e desvantagem do classificador por redes neurais disponível no *software ENVI* em relação àquele implementado no ambiente *MATLAB*, é que o disponível no *ENVI* não permite a manipulação do usuário quanto à arquitetura da rede, ou seja, não é possível inserir mais que duas camadas de neurônios escondidos, muito menos permite a variação de números de neurônios inseridos nestas camadas, o que faz com que cause uma certa deficiência no resultado. Fica muito mais difícil conseguir “achar” a arquitetura ideal, o que já não acontece em relação ao sistema disponível no ambiente *MATLAB*, que permite que estes parâmetros seja manipulado pelo usuário.

As amostras coletadas para a classificação tanto para o método das redes neurais, quanto para o da máxima verossimilhança foram as mesmas, porém a diferença foi em relação ao número de *pixels* coletados para cada uma delas, já que para a classificação através da máxima verossimilhança, há a necessidade de que o número de *pixels* seja suficientemente grande para que as estimativas dos parâmetros populacionais das classes sejam adequadas.

Para a classificação pelo método das redes neurais quanto menor for o número de *pixels* para cada amostra, melhor é o resultado da classificação. Na classificação pelo método das redes neurais implementada no *software ENVI*, há a necessidade da escolha do número de camadas escondidas e de iterações de parada. Através de um processo difícil e trabalhoso de *tentativa e erro*, o melhor resultado obtido foi utilizando apenas 1 camada escondida, em 2000 iterações.

Para a análise dos resultados foi utilizada a ferramenta do próprio *software*, que dispõe do cálculo da *matriz de confusão*, onde apresenta a acurácia através do *coeficiente de concordância κ* .

A comparação dos resultados entre estes métodos foi realizada através deste índice, como também da análise visual entre os mapas temáticos resultantes da classificação.

3.2.7 COMPARAÇÃO DA CLASSIFICAÇÃO ORIENTADA AO OBJETO POR REDES NEURAIS ARTIFICIAIS E PELO MÉTODO DA DISTÂNCIA MÍNIMA DO SOFTWARE ECOGNITION.

O *software eCognition* permite a análise de imagens em relação a *objetos* homogêneos extraídos da mesma, levando em consideração a sua forma, as relações espaciais e hierárquicas entre os mesmos. Para a classificação baseada na metodologia de dados *orientados ao objeto*, o *software* dispõe de dois classificadores, o *vizinho mais próximo* e a *lógica Fuzzy*. Ambos consideram as características de *objeto* e não de *pixels isolados*. No presente trabalho, utilizou-se do *vizinho mais próximo*, já que o classificador através da *lógica fuzzy*, necessita do conhecimento do usuário a respeito das características de cada tipo de objeto, tarefa a qual nem sempre é fácil, pois é difícil propor funções de pertinência, mesmo para fatos bem conhecidos numa análise visual. Optou-se por utilizar o método do *vizinho mais próximo*, pois ele não requer

um alto conhecimento das funções de pertinência, diminuindo com isto a participação do usuário, de maneira similar ao que ocorre na classificação por redes neurais.

O aplicativo *ECOGNITION* permite que através da imagem segmentada na tela (agregada em regiões), sejam coletadas amostras para descrever cada classe a partir das regiões. As amostras utilizadas foram as mesmas utilizadas na classificação por redes neurais. Com a finalidade de tornar a comparação mais justa, foram utilizados, além das propriedades espectrais, alguns parâmetros espaciais dos segmentos, entre os quais estão: a *área*, *perímetro*, *comprimento*, *largura*, *razão entre comprimento e largura e direção principal* (que corresponde à direção do auto-vetor correspondente ao maior autovalor). Procurou-se utilizar os índices de forma mais próximos aos utilizados na classificação no *MATLAB*.

3.2.8 CLASSIFICAÇÃO DE OUTRA ÁREA DA IMAGEM IKONOS DE MATINHOS.

Com o objetivo de se observar se os parâmetros encontrados para a classificação pelo método das redes neurais, poderiam ser utilizados para a classificação de uma outra área, fez-se o recorte da imagem cujas classes observadas foram:

- (1) construções de amianto;
- (2) construções brancas;
- (3) construções de cor verde;
- (4) ruas de asfalto (e Rodoviária);
- (5) areia (ruas e solo com areia);
- (6) vegetação de alta densidade;
- (7) solo exposto, e;
- (8) sombra.

Como se pôde observar, as classes observadas foram praticamente as mesmas,

com uma única diferença de a segunda imagem (Figura 24), não apresentar áreas com vegetação de alta densidade.

FIGURA 24 – SEGUNDA IMAGEM DE MATINHOS UTILIZADA NA CLASSIFICAÇÃO POR REDES NEURAIS ARTIFICIAIS.



Através da análise da Figura 24, se pode perceber que há uma maior concentração da área urbana em relação a outra área utilizada, como também se observa ainda a presença de construções com a mesma resposta espectral das ruas de areia, e construções em amianto com ruas de asfalto.

Nesta fase, foi realizada a classificação utilizando o *software E-COGNITION*, a fim de comparação, onde se utilizou o mesmo método e parâmetros usados na classificação da primeira imagem. Para esta análise, a classificação por redes neurais foi realizada no ambiente *MATLAB* da mesma forma que para a imagem anterior, porém foi feita apenas a classificação utilizando apenas os parâmetros espectrais (*IVP-R-G*) e a melhor combinação de parâmetros encontrada para a imagem anterior (*IVP-R-G + compacidade + razão eixos + 1º momento*).

A classificação foi realizada utilizando os parâmetros de duas maneiras:

(1) utilizando os pesos encontrados para a classificação realizada anteriormente para a outra imagem estudada. Estes pesos foram os encontrados no treinamento da rede para a outra imagem;

(2) treinando outra rede para esta imagem, porém utilizando o mesmo número de neurônios na camada escondida (138, 46) e utilizando 15 amostras para cada classe.

CAPÍTULO 4

RESULTADOS E DISCUSSÕES

4.1 SEGMENTAÇÃO DA IMAGEM

Para a obtenção da segmentação da imagem utilizada neste trabalho, a fim da extração dos objetos utilizados na posterior classificação, optou-se pela utilização do *Software eCognition*, que faz a análise de dados orientada *a objeto*. Ele permite a extração de objetos homogêneos da imagem na resolução desejada, e descrevê-los levando em consideração características espectrais e de forma.

Para a segmentação neste *software*, há a necessidade da definição de vários parâmetros, entre os quais os principais são: *scale parameter* e *homogeneity criterion*. Realizaram-se vários testes com a variação destes parâmetros, principalmente em relação ao parâmetro de escala, a fim de se obter um tamanho ideal para os objetos. Esta fase é considerada muito importante, já que o resultado da mesma é a base para a posterior classificação *orientada ao objeto*.

Considerou-se que a segmentação considerada ideal para a imagem IKONOS utilizada neste trabalho deveria estar composta por segmentos alongados para a representação das ruas, e segmentos regulares para a representação das construções. Os segmentos devem também ter alta uniformidade espacial.

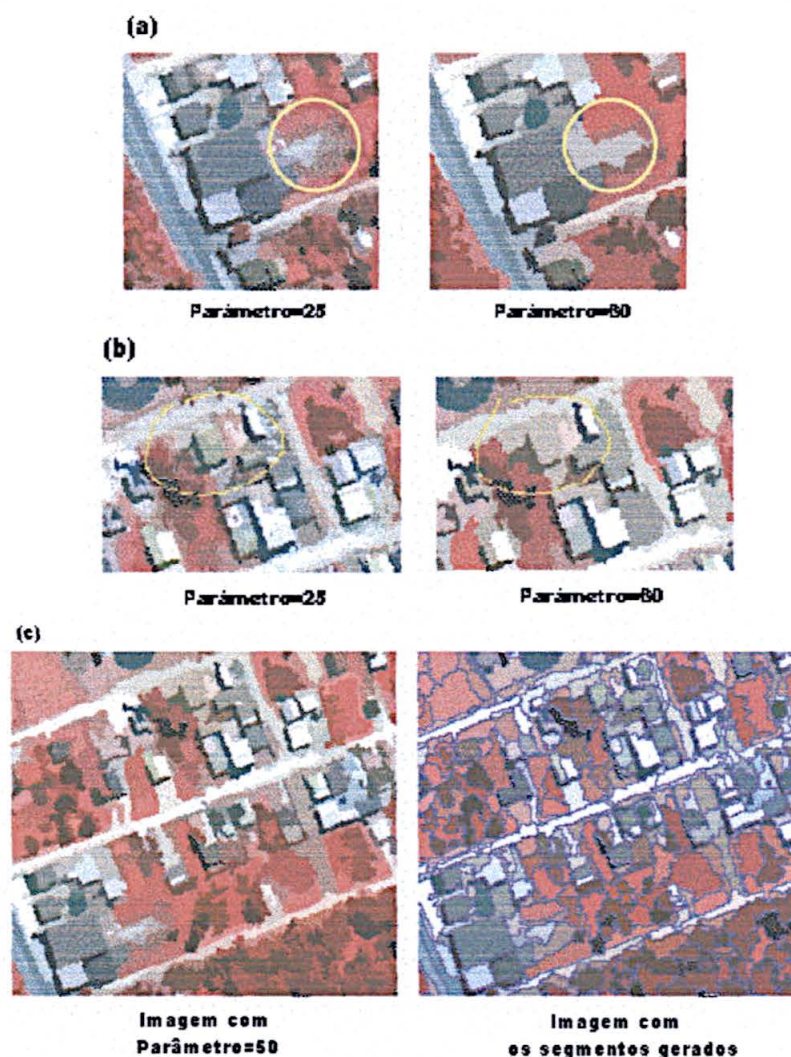
Portanto, a determinação dos parâmetros ideais é bastante complicada. Ocorre que um certo parâmetro de escala pode ser ótimo para segmentar uma classe e ser, ao mesmo tempo, muito deficiente para separar outros objetos. Procurou-se encontrar um parâmetro intermediário que fosse ideal para todos os padrões da imagem.

Pode-se observar através da Figura 25 (a) que, para o parâmetro de escala “25”, as construções ficaram muito fragmentadas, assim como as ruas. Já para um parâmetro

de escala maior, por exemplo 60, os segmentos ficaram “grandes”. O efeito de um parâmetro muito grande é que regiões com objetos diferentes podem ser fundidas para formar uma região maior. Isto pode ser visto na Figura 25 (b), onde algumas construções e regiões ocupadas por areia foram fundidas.

Através dos testes, verificou-se que o parâmetro de escala ideal seria em torno de “50”, como é mostrado na Figura 25 (c). Este parâmetro resulta da combinação de dois critérios de uniformidade: espacial e espectral. Para gerar objetos mais compactos, optou-se por escolher pesos iguais para ambos (50% como peso para cor e 50% para a forma).

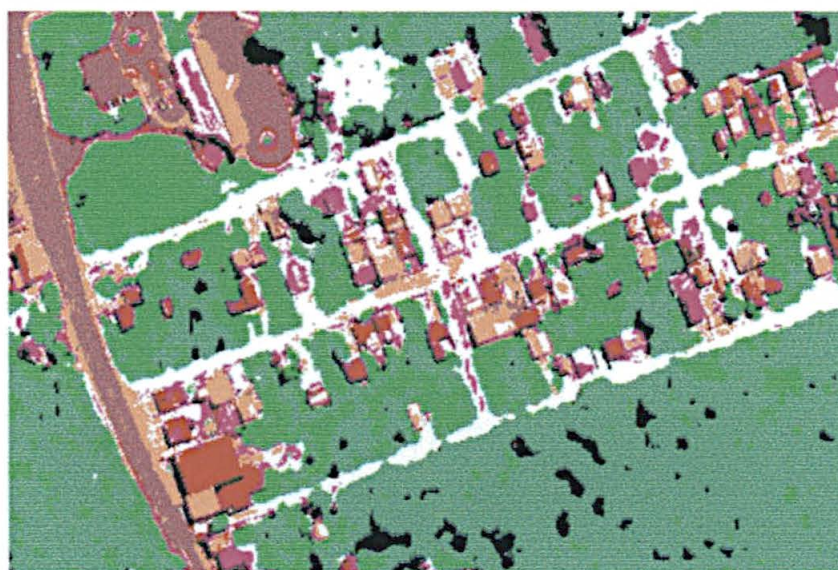
FIGURA 25 – IMAGEM SEGMENTADA EM PARÂMETROS DE ESCALA DIFERENTES. (A) IMAGEM SEGMENTADA COM PARÂMETRO DE ESCALA DE 25. (B) IMAGEM SEGMENTADA COM PARÂMETRO DE ESCALA DE 60. (C) IMAGEM SEGMENTADA COM PARÂMETRO DE ESCALA DE 50.



4.2 CLASSIFICAÇÃO *PIXEL A PIXEL* POR REDES NEURAIIS E MÉTODO DA MÁXIMA VEROSSIMILHANÇA

As Figuras 26 e 27, respectivamente apresentam o resultado obtido na classificação utilizando os métodos máxima verossimilhança ($\kappa= 0.9226$), e redes neurais ($\kappa= 0.9126$). Ambas foram realizadas com o auxílio do *software ENVI*.

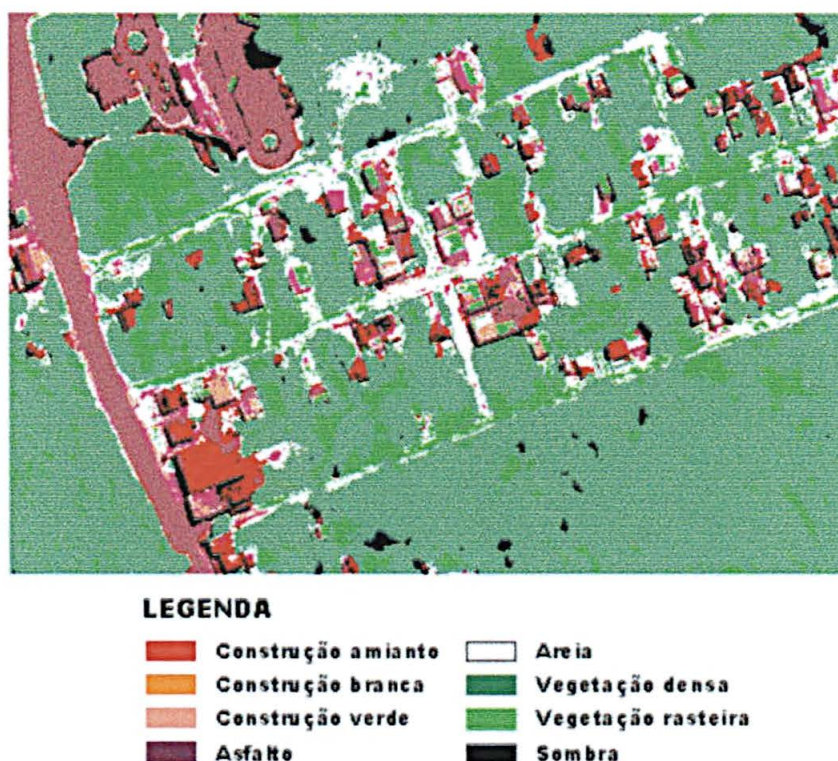
FIGURA 26 – RESULTADO DA CLASSIFICAÇÃO UTILIZANDO O MÉTODO DA MÁXIMA VEROSSIMILHANÇA (ENVI).



LEGENDA

 Construção amianto	 Areia
 Construção branca	 Vegetação densa
 Construção verde	 Vegetação rasteira
 Asfalto	 Sombra

FIGURA 27 – RESULTADO DA CLASSIFICAÇÃO UTILIZANDO O MÉTODO DAS REDES NEURAIS ARTIFICIAIS (ENVI).



Apesar do excelente resultado de ambas as classificações, praticamente o mesmo resultado apresentado para o *coeficiente κ* , se pode perceber uma certa confusão entre as classes *construções* e *areia* e até mesmo entre a classe *asfalto*. Considerando a dificuldade apresentada quanto à definição da arquitetura no *software*, nota-se uma que os padrões apresentaram-se de uma forma mais homogênea (ruas de asfalto, por exemplo) em comparação ao método da máxima verossimilhança.

As tabelas 8 e 9, mostram a matriz de confusão para a classificação por redes neurais e pelo método da máxima verossimilhança. Analisando a Tabela 9, nota-se que na classificação pelo método da máxima verossimilhança houve muito mais confusão entre as classes *construções* e *ruas de asfalto*. Este fato aponta uma vantagem do método das redes neurais. Pode-se concluir, ainda, que se houvesse possibilidade de definir uma melhor arquitetura para a rede, os resultados poderiam ser ainda melhores.

Se pode observar que o resultado da classificação por redes neurais o resultado para a classe *vegetação* ficou mais homogênea, como também as ruas de asfalto.

A classificação através da Máxima Verossimilhança apresentou um melhor resultado para as ruas de areia, e para as construções as redes apresentaram um melhor resultado, já que se observa que houve uma melhor discriminação entre as mesmas e a malha viária.

Através da análise das Tabelas 8 e 9, pode-se perceber que a classificação por Máxima Verossimilhança apresenta um número maior de *pixels* de treinamento na geração do resultado em comparação às Redes Neurais. Esta diferença se faz necessária pois, para que o resultado por redes neurais seja satisfatório, há a necessidade de um número mínimo de amostras para o treinamento, ao contrário do método da máxima verossimilhança, que necessita de um número razoável de amostras para o posterior cálculo estatístico.

As tabelas apresentam também os valores da acurácia global, que também representa uma forma de análise da precisão, e a mesma é obtida através da razão entre o somatório dos elementos da diagonal da matriz, pelo número total de elementos (*pixels*) coletados como amostras de treinamento.

TABELA 8 – MATRIZ DE CONFUSÃO RESULTANTE DA CLASSIFICAÇÃO ATRAVÉS DO MÉTODO DAS REDES NEURAS ARTIFICIAIS (ENVI).

Acurácia Global		92,35%								
Coeficiente		91,26								
Kappa		91,26								
Class	Ground Truth (pixel)									
	amianto	branca	verde	asfalto	areia	areia	veget densa	veget rast	sombra	Total
const amianto	24									24
branca		8								8
verde			16							16
asfalto				20						20
areia		4			20					24
areia						2				2
veget densa							30			30
veget rast						10		25		35
sombra									24	24
Total	24	12	16	20	20	12	30	25	24	183

TABELA 9 – MATRIZ DE CONFUSÃO RESULTANTE DA CLASSIFICAÇÃO ATRAVÉS DO MÉTODO DA MÁXIMA VEROSSIMILHANÇA (ENVI).

Acurácia global	93,46%								
Coefficiente Kappa	92,26%								
Class	Ground Truth (pixel)								
	amianto	branca	verde	asfalto	areia	veget densa	veget rast	sombra	Total
const amianto	180								180
branca		162		5	55				222
verde	2	7	159		12				180
asfalto				315					315
areia		17	8		196				221
veget densa						507	7	1	515
veget rast						20	337		357
sombra	1							75	76
Total	183	186	167	320	263	527	344	76	2066

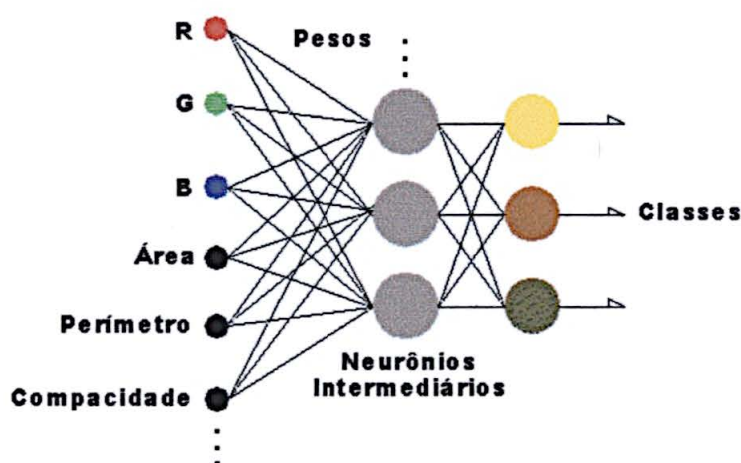
4.3 CLASSIFICAÇÃO ORIENTADA AO OBJETO UTILIZANDO REDES NEURAIS ARTIFICIAIS

Coletadas as amostras de treinamento e com ajuda dos programas auxiliares, procedeu-se à montagem da tabela de entrada das amostras de treinamento para a rede. Iniciou-se uma série de testes, sendo estes realizados através da análise dos parâmetros de forma inseridos às entradas de acordo com a Tabela 10.

A classificação por redes neurais utilizando o *toolbox* do ambiente *MATLAB* é considerada bastante simples, e possui a vantagem de se poder manipular o tipo de rede a ser implantada, o número de neurônios das camadas escondidas e o número destas camadas, e a manipulação de outros parâmetros, tais como: o número de iterações a serem realizadas para a convergência e o erro de aprendizado. Por outro lado, o ambiente possui um tempo de treinamento razoável (variando de 3min a 15min, isto porque com a inserção dos parâmetros o tempo de processamento aumenta), em comparação ao *software eCognition* (em torno de 30s).

A classificação foi realizada de acordo com a Figura 28, onde a camada de entrada da rede foi composta em função do número das entradas em teste, as camadas escondidas foram sempre utilizadas da mesma maneira para todos os testes, ou seja, duas camadas de neurônios de 138 e 46 respectivamente, e a última camada foi representada binariamente, ou seja com apenas três neurônios, representando até 8 classes.

FIGURA 28 – ESQUEMA DA ARQUITETURA DA REDE UTILIZADA.



Iniciou-se pela classificação puramente espectral, com o objetivo de comparar seu resultado com a classificação pelo método da máxima verossimilhança *pixel a pixel* e com o resultado da aplicação das redes neurais *pixel a pixel*. O resultado também foi utilizado para estimar o ganho que é possível obter a com a inserção dos índices de forma.

Em princípio os testes foram iniciados inserindo um a um os índices de forma às propriedades espectrais, e se pode observar através da Tabela 10, que os melhores resultados encontrados para os índices de forma foram: *área*, *compacidade* e *5º momento*.

TABELA 10 – RESULTADO DAS CLASSIFICAÇÕES REALIZADAS COM OS ÍNDICES DE FORMA INSERIDOS UM A UM.

PARÂMETROS UTILIZADOS	ACURÁCIA GLOBAL
APENAS IVP-R-G	0.6105
IVP-R-G + ÁREA	0.7368
IVP-R-G + PERÍMETRO	0.6842
IVP-R-G + DEFIC. CONVEXA	0.6316
IVP-R-G + COMPACIDADE	0.7789
IVP-R-G + RAZÃO EIXOS	0.6526
IVP-R-G + 1º MOMENTO	0.6526
IVP-R-G + 2º MOMENTO	0.6842
IVP-R-G + 3º MOMENTO	0.6316
IVP-R-G + 4º MOMENTO	0.6737
IVP-R-G + 5º MOMENTO	0.7263
IVP-R-G + 6º MOMENTO	0.6000
IVP-R-G + 7º MOMENTO	0.6526
IVP-R-G + CÓDIGO DA CADEIA	0.6316

Procedeu-se então a avaliar o efeito da combinação de duas variáveis de cada vez. Ou seja, foram inseridos, além dos parâmetros espectrais, dois índices de forma na classificação, e o resultado pode ser observado através da Tabela 11. Os melhores resultados obtidos foram para as combinações dos parâmetros: (*compacidade e razão eixos*); (*compacidade e 1º momento*).

TABELA 11 – RESULTADO DAS CLASSIFICAÇÕES REALIZADAS COM OS ÍNDICES DE FORMA INSERIDOS DOIS A DOIS.

PARÂMETROS UTILIZADOS	ACURÁCIA GLOBAL
IVP-R-G + ÁREA + COMPACIDADE	0.6632
IVP-R-G + PERÍMETRO + COMPACIDADE	0.6316
IVP-R-G + PERÍMETRO + RAZÃO EIXOS	0.6632
IVP-R-G + ÁREA + RAZÃO EIXOS	0.6316
IVP-R-G + COMPACIDADE + RAZÃO EIXOS	0.7474
IVP-R-G + COMPACIDADE + 1º MOMENTO	0.7474

Para os resultados com três índices de forma, as melhores classificações foram com as combinações dos parâmetros, Tabela 12:

- *compacidade + razão eixos + 1º momento (0.7789);*
- *compacidade + razão eixos + deficiência convexa (0.7684);*
- *compacidade + 1º e 2º momentos (0.7474);*
- *compacidade + razão eixos + código da cadeia (0.7263), e;*
- *perímetro + razão eixos + código da cadeia (0.7053).*

TABELA 12 – RESULTADO DAS CLASSIFICAÇÕES REALIZADAS COM TRÊS ÍNDICES DE FORMA ADICIONADAS AOS PARÂMETROS ESPECTRAIS.

PARÂMETROS UTILIZADOS	ACURÁCIA GLOBAL
IVP-R-G + ÁREA + PERÍMETRO + COMPACIDADE	0.6737
IVP-R-G + ÁREA + PERÍMETRO + RAZÃO EIXOS	0.6947
IVP-R-G + ÁREA + COMPACIDADE + RAZÃO EIXOS	0.6316
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS	0.6842
IVP-R-G + PERÍMETRO + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA	0.6632
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + DEFIC. CONVEXA	0.7684
IVP-R-G + PERÍMETRO + RAZÃO EIXOS + CÓDIGO DA CADEIA	0.7053
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + CÓDIGO DA CADEIA	0.7263
IVP-R-G + PERÍMETRO + COMPACIDADE + DEFICIÊNCIA CONVEXA	0.6737
IVP-R-G + PERÍMETRO + COMPACIDADE + CÓDIGO DA CADEIA	0.6105
IVP-R-G + PERÍMETRO + COMPACIDADE + 1º MOMENTO	0.6632
IVP-R-G + PERÍMETRO + RAZÃO EIXOS + 1º MOMENTO	0.6316
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º MOMENTO	0.7789
IVP-R-G + COMPACIDADE + 1º E 2º MOMENTOS	0.7474

Foram então sendo aumentadas as combinações das entradas de dados para a classificação a fim de se avaliar o grau de relevância neste aumento. Para este grupo de entradas foram inseridos aos parâmetros espectrais mais quatro parâmetros da representação da forma (Tabela 13). O melhor resultado encontrado foi para a combinação: *perímetro + compacidade + razão eixos + 1º momento*.

TABELA 13 – RESULTADO DAS CLASSIFICAÇÕES REALIZADAS COM QUATRO ÍNDICES DE FORMA ADICIONADAS AOS PARÂMETROS ESPECTRAIS.

PARÂMETROS UTILIZADOS	ACURÁCIA GLOBAL
IVP-R-G + ÁREA + PERÍMETRO + COMPAC. + RAZÃO EIXOS	0.6211
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA	0.6947
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + CÓDIGO DA CADEIA	0.6632
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + 1º MOMENTO	0.7053
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA + 1º MOMENTO	0.6842
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º E 2º MOMENTOS	0.6947
IVP-R-G + PERÍMETRO + COMPACIDADE + 1º E 2º MOMENTOS	0.6737
IVP-R-G + PERÍMETRO + RAZÃO EIXOS + 1º E 2º MOMENTOS	0.6842

Para o restante dos testes realizados foram sendo considerados somente os índices de maior relevância, e então testados as melhores combinações entre eles. As combinações com cinco índices como os seus respectivos valores de acurácia são apresentados na Tabela 14. Os melhores resultados encontrados para este grupo foram:

- *compacidade + razão eixos + 1º, 2º e 3º momentos (0.7579);*
- *perímetro + compacidade + razão eixos + código da cadeia + 1º momento (0.7474).*

TABELA 14 – RESULTADO DAS CLASSIFICAÇÕES REALIZADAS COM CINCO ÍNDICES DE FORMA ADICIONADAS AOS PARÂMETROS ESPECTRAIS.

PARÂMETROS UTILIZADOS	ACURÁCIA GLOBAL
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA + CÓDIGO DA CADEIA	0.6737
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + CÓDIGO DA CADEIA + 1º MOMENTO	0.7474
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + 1º E 2º MOMENTOS	0.6316
IVP-R-G + PERÍMETRO + COMPACIDADE + 1º, 2º E 3º MOMENTOS	0.6737
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º E 3º MOMENTOS	0.7579

O mesmo procedimento foi realizado para as demais combinações de 6 a 9

parâmetros de forma e os resultados são apresentados na Tabela 15. Os melhores resultados encontrados foram:

- *compacidade + todos os 7 momentos (0.7684);*
- *compacidade + razão eixos + 1º, 2º e 3º momentos (0.7579);*
- *compacidade + razão eixos + 1º ao 6º momentos (0.7579);*
- *compacidade + razão eixos + todos os 7 momentos (0.7579), e;*
- *perímetro + compacidade + todos os 7 momentos (0.7368).*

TABELA 15 – RESULTADO DAS CLASSIFICAÇÕES REALIZADAS COM SEIS A NOVE ÍNDICES DE FORMA ADICIONADAS AOS PARÂMETROS ESPECTRAIS.

PARÂMETROS UTILIZADOS	ACURÁCIA GLOBAL
IVP-R-G + PERÍMETRO + COMPACIDADE + RAZÃO EIXOS + DEFICIÊNCIA CONVEXA + CÓDIGO DA CADEIA + 1º MOMENTO	0.6421
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º E 3º MOMENTOS	0.7579
IVP-R-G + TODOS OS 7 MOMENTOS	0.6947
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º, 3º, 4º E 5º MOMENTOS	0.6947
IVP-R-G + COMPACIDADE + TODOS OS 7 MOMENTOS	0.7684
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + 1º, 2º, 3º, 4º, 5º E 6º MOMENTOS	0.7579
IVP-R-G + COMPACIDADE + RAZÃO EIXOS + TODOS OS 7 MOMENTOS	0.7579
IVP-R-G + PERÍMETRO + COMPACIDADE + TODOS OS 7 MOMENTOS	0.7368

O auxílio dos parâmetros de forma utilizados como dado auxiliar na classificação resultou em um ganho de 17% no valor da acurácia, Figura 29.

FIGURA 29 – RESULTADO DAS CLASSIFICAÇÕES PURAMENTE ESPECTRAL (IVP-R-G) E COM OS ÍNDICES DE FORMA COMPACIDADE + RAZÃO EIXOS + 1º MOMENTO.



4.4 CLASSIFICAÇÃO ORIENTADA A OBJETO POR REDES NEURAIIS ARTIFICIAIS E ATRAVÉS DO *SOFTWARE ECOGNITION*.

No presente trabalho, utilizou-se do classificador *vizinho mais próximo*, já que o classificador através da lógica *fuzzy* necessita do conhecimento do usuário a respeito das características de cada tipo de objeto, tarefa a qual nem sempre é fácil, pois é difícil propor funções de pertinência, mesmo para fatos bem conhecidos numa análise visual.

Procurou-se utilizar os índices de forma mais próximos aos utilizados na classificação no *MATLAB*. Os parâmetros utilizados para representar a forma foram: a *área*, *perímetro*, *comprimento*, *largura*, *razão entre comprimento e largura* e *direção principal* (que corresponde à direção do auto-vetor correspondente ao maior autovalor).

As Figuras 30 e 31 apresentam os resultados da classificação realizada através do *ECOGNITION* e o da melhor classificação obtida através das redes neurais, onde os parâmetros utilizados foram *compacidade*, *razão eixos* e *1º momento*.

Comparando ambas as classificações com a imagem original mostrada na Figura 32, é possível se observar claramente que a classificação por redes neurais obteve um resultado superior à realizada através do *software ECOGNITION*. Vale a pena ressaltar que o método de classificação usado no *ECOGNITION* é o método de classificação mais simples e que provavelmente aplicando a classificação baseada em funções de pertinência definidas pelo usuário este resultado poderia ser melhor. Porém, o que se pretendeu comparar é a eficiência de processos de classificação onde a participação do usuário é mínima e avaliar mostrar a maior facilidade da rede em generalizar entradas sem que se necessite de um alto grau de conhecimento tanto das características dos objetos da área de trabalho, quanto da manipulação das funções de pertinência necessárias a este método de classificação.

Através da análise das figuras, se pode perceber a maior homogeneidade entre as classes no resultado realizado pelas redes neurais, a vegetação rasteira está muito mais homogênea, e na classificação feita no *software ECOGNITION* houve muito mais confusão entre as classes *construções* e *areia*.

FIGURA 30 – CLASSIFICAÇÃO REALIZADA NO SOFTWARE ECOGNITION.

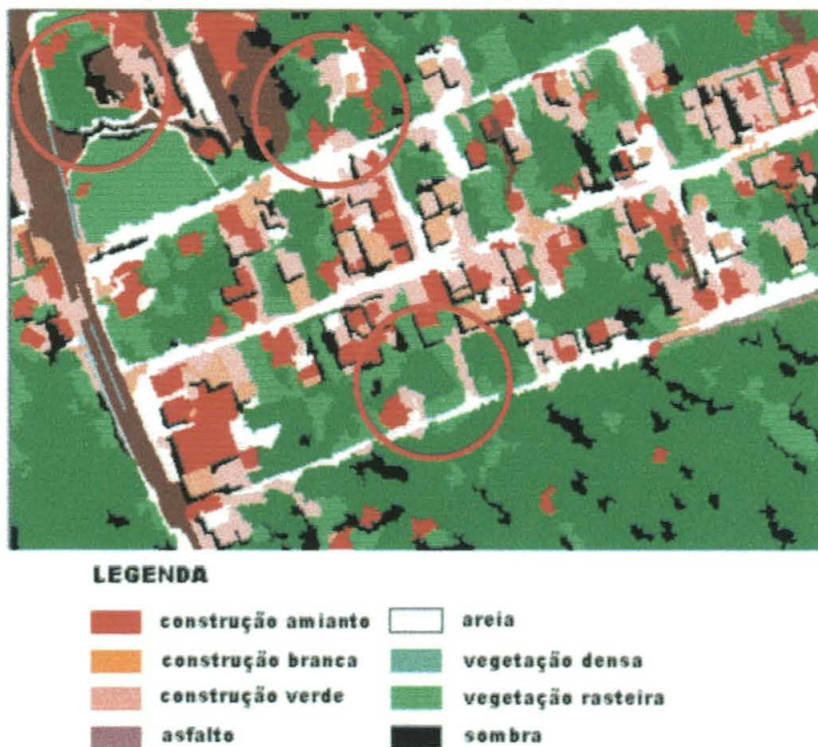


FIGURA 31 – CLASSIFICAÇÃO REALIZADA PELO MÉTODO DAS REDES NEURAIS ARTIFICIAIS.



FIGURA 32 – IMAGEM ORIGINAL.



4.5 VERIFICAÇÃO DA EFICIÊNCIA DA REDE NEURAL EM OUTRA ÁREA

Utilizou-se uma outra região da imagem de Matinhos com o intuito de se observar se os parâmetros encontrados para a classificação pelo método das redes neurais poderiam ser utilizados para a classificação de uma região de estudo maior, ou seja, se o método é válido para outros conjuntos de *pixels*.

Um novo recorte com as mesmas classes foi selecionado e submetido a três classificações. Este recorte foi primeiramente segmentado, da mesma forma que nos experimentos anteriores. Com a finalidade de avaliar se a rede teria um bom desempenho nesta região, duas classificações foram feitas: Na primeira, o recorte foi classificado pelo método das redes neurais, usando apenas as variáveis espectrais. No segundo, a região foi classificada usando as variáveis espectrais e os parâmetros de forma: *compacidade, razão eixos e 1º momento*.

A comparação dos resultados destas classificações foi realizada apenas visualmente. Comparando a imagem original (Figura 33), com as imagens resultante

da classificação puramente espectral (Figuras 36), nota-se que a classificação pelas redes apresenta ainda certa confusão entre as classes *construções e areia*.

Já para a classificação utilizando um novo treinamento (Figura 37) com o mesmo número de neurônios nas camadas escondidas (138, 46), como também utilizando 15 amostras de treinamento para cada classe, utilizando os mesmos parâmetros de forma *compacidade, razão eixos e 1º momento*, apresentou uma acurácia de 0,9375. Este resultado apresentou uma melhora de 6% em relação à classificação puramente espectral, apresentada na Figura 36.

Comparando estes resultados com o resultado que poderia ser obtido usando o algoritmo de classificação por distância mínima do *ECOGNITION* (Figura 34), nota-se que o desempenho da rede neural é superior em ambos casos.

FIGURA 33 – IMAGEM ORIGINAL.



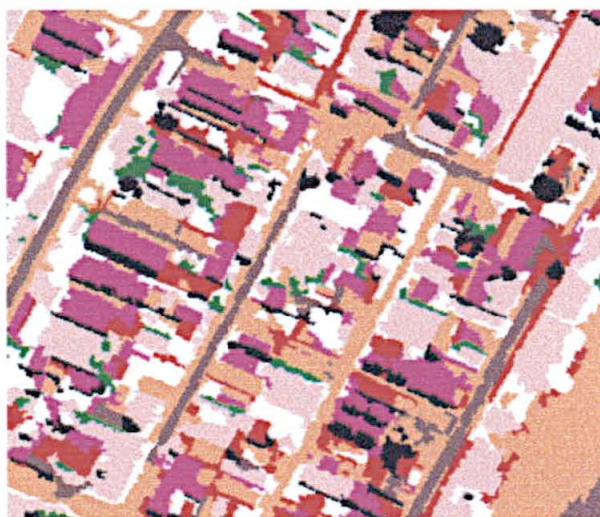
FIGURA 34 – IMAGEM RESULTANTE DA CLASSIFICAÇÃO UTILIZANDO PARÂMETROS ESPECTRAIS E FORMA, ATRAVÉS DO SOFTWARE ECOGNITION.



LEGENDA

construção amianto	areia
construção branca	vegetação densa
construção verde	vegetação rasteira
asfalto	sombra

FIGURA 35 – IMAGEM RESULTANTE DA CLASSIFICAÇÃO UTILIZANDO OS MESMOS PARÂMETROS DE ARQUITETURA DA REDE DA IMAGEM UTILIZADA ANTERIORMENTE.



LEGENDA

construção amianto	areia
construção branca	vegetação densa
construção verde	vegetação rasteira
asfalto	sombra

FIGURA 36 – IMAGEM RESULTANTE DA CLASSIFICAÇÃO PURAMENTE ESPECTRAL (IVP-R-G), UTILIZANDO AS REDES NEURAIS ARTIFICIAIS.



O resultado obtido pela classificação utilizando os mesmos parâmetros treinados para a rede da imagem anterior (Figura 35), onde se utilizou os pesos usados no treinamento, não apresentou um resultado satisfatório e a acurácia global encontrada foi de 0,5500.

Já para a classificação utilizando um novo treinamento (Figura 37), com o mesmo número de neurônios nas camadas escondidas (138, 46), como também utilizando 15 amostras de treinamento para cada classe, utilizando os mesmos parâmetros de forma *compacidade*, *razão eixos* e *1º momento*, apresentou uma acurácia de 0,9375. Este resultado apresentou uma melhora de 6% em relação à classificação puramente espectral, apresentada na Figura 36. Isto pode ser observado em algumas regiões onde havia uma certa confusão entre construções com o mesmo valor de reflectância de ruas de areia. Com a inserção dos parâmetros descritores da forma, houve a separação destas classes.

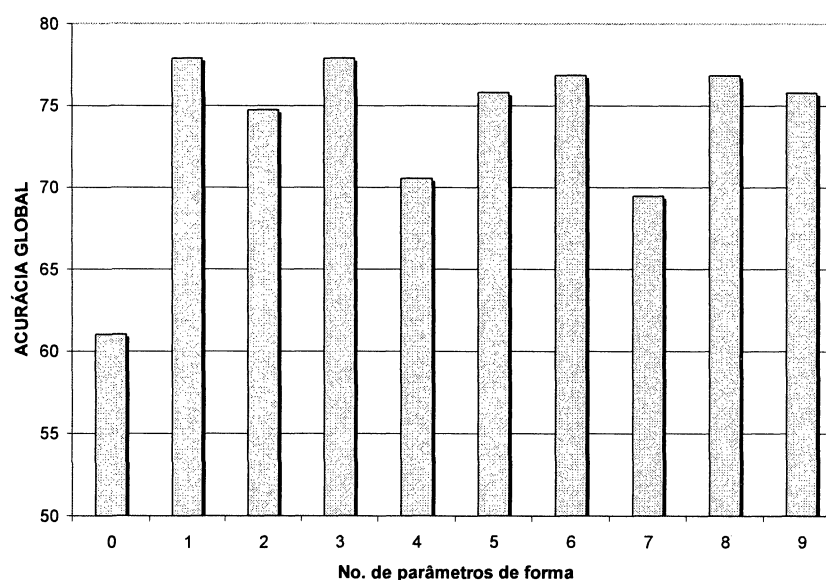
FIGURA 37 – IMAGEM RESULTANTE DA CLASSIFICAÇÃO COM NOVO TREINAMENTO PARA A REDE, UTILIZANDO A MELHOR COMBINAÇÃO DOS PARÂMETROS DE FORMA ENCONTRADOS PARA A IMAGEM ANTERIOR.



4.6 ESCOLHA DO MELHOR PARÂMETRO DE FORMA

Os testes realizados comprovam que a adição de informação espacial no processo de classificação pode contribuir significativamente para a melhora dos resultados. No entanto, torna-se difícil escolher o melhor parâmetro de forma para a solução do problema, visto que existe uma grande variedade de parâmetros de forma citados na literatura. O sucesso obtido na série de testes pode ser avaliado através da Figura 38. Nela, os melhores resultados são representados em função do número de parâmetros de forma utilizados.

FIGURA 38 – ACURÁCIA GLOBAL EM FUNÇÃO DO NÚMERO DE PARÂMETROS DE FORMA.



Observa-se que um ganho significativo é verificado com a inclusão de pelo menos uma variável de forma. Neste caso, trata-se do índice de compacidade. Na medida em que mais variáveis são incluídas, a acurácia global tende a diminuir, provavelmente porque a arquitetura da rede neural tem problemas para selecionar a variável mais significativa e com isto gera resultados piores. No entanto, qualquer

combinação testada gerou resultados melhores que uma classificação puramente espectral.

O teste realizado utilizando uma área diferente também comprova esta afirmação, embora o ganho seja menor. No entanto, deve-se considerar que, neste caso, o classificador puramente espectral já tinha atingido um desempenho muito bom.

CAPÍTULO 5

CONCLUSÕES E RECOMENDAÇÕES

No desenvolvimento desta pesquisa, foram apresentados diferentes experimentos com o objetivo de avaliar a eficiência da inclusão de informação a respeito da forma na classificação de imagens de satélite de alta resolução. Através da análise dos resultados obtidos com a classificação utilizando parâmetros de forma como um dado auxiliar como também a utilização das redes neurais artificiais, pode-se concluir que os classificadores tradicionais *pixel-a-pixel*, por apenas utilizarem as informações espectrais, mostram limitações quando aplicados a imagens de alta resolução espacial. Com a inserção dos novos sensores de alta resolução, onde maiores detalhes da paisagem urbana podem ser identificados, os métodos de classificação orientados ao objeto passam a ter maior importância, já que os mesmos consideram além dos atributos espectrais as relações espaciais entre os *pixels*.

A inclusão de parâmetros auxiliares, como a forma ou a textura, constitui um fator muito relevante para o aprimoramento da classificação dos objetos presentes na superfície da Terra, principalmente para imagens onde há muitos padrões distintos com a mesma aparência espectral.

Como a classificação *orientada a objeto* é dependente de uma prévia segmentação da imagem, a mesma deve ser a mais rigorosa possível. Porém ainda há uma certa dificuldade em se obter uma segmentação ideal para todos os padrões da imagem, talvez em função de uma carência por parte dos algoritmos de segmentação. O que se pôde perceber foi que o resultado da classificação ficou limitado ao resultado encontrado pela segmentação, ou seja, a acurácia da classificação é totalmente dependente do resultado da segmentação. Erros já vindos da segmentação, não poderão ser corrigidos na classificação.

Apesar do resultado da classificação *orientada ao objeto* estarem condicionados ao sucesso da segmentação, a mesma apresenta uma acurácia superior à classificação *pixel a pixel*, e que o mapa gerado apresenta uma maior homogeneidade entre as classes, o que facilita a interpretação do mapa, principalmente se o mesmo for utilizado em um sistema de informações geográficas - *SIG*, onde poderão ser extraídos valores quantitativos. A inclusão da forma é uma ferramenta útil para separar classes espectralmente similares, que ocorrem freqüentemente em imagens de alta resolução espacial de áreas urbanas. A forma, no entanto, não é facilmente descrita utilizando parâmetros derivados de imagens em formato “raster”. Por este motivo existem diferentes parâmetros e pode ocorrer que alguns sejam úteis para descrever determinadas classes e fracassem na descrição de outras. Portanto, a utilização de vários parâmetros, de maneira conjunta, parece ser a melhor opção.

A inclusão de parâmetros de forma, principalmente os considerados mais relevantes neste trabalho, mostrou-se satisfatória na detecção da malha viária e construções nas imagens de alta resolução com características espectrais similares.

Para imagens com as mesmas características, ou seja, que apresentem praticamente as mesmas classes, os parâmetros da arquitetura da rede permanecem os mesmos, porém quando não se têm coincidência de todas as classes há a necessidade de um novo treinamento da rede, não podendo ser utilizados os pesos usados na imagem treinada anteriormente.

Quanto à utilização das redes neurais artificiais, esta se mostrou muito vantajosa na tarefa de integrar dados de naturezas diferentes na classificação, como é a informação espectral e os parâmetros espaciais, já que se podem obter resultados satisfatórios utilizando-se de um número mínimo de amostras para que a mesma faça a generalização. A flexibilidade dos algoritmos de classificação baseados em redes neurais permite integrar variáveis diferentes na identificação de objetos, como por exemplo de características espectrais e espaciais. Além da vantagem de não haver a

necessidade prévia de um grande conhecimento do usuário quanto à área e a formulação de regras de decisão entre os padrões existentes.

Os melhores resultados obtidos para a classificação utilizando a forma foram para as variáveis (*IVP-R-G + compacidade*) e (*IVP-R-G + compacidade + razão eixos + 1º momento*). Se pode então concluir que para a área estudada, a compacidade foi muito relevante seguida da razão entre os eixos principais e o 1º momento, isto porque as formas dos segmentos que geraram maior confusão foram as mais próximas de um quadrado (relativas às construções) com as mais alongadas (relativas às ruas). Utilizando apenas a compacidade, o resultado foi muito satisfatório, isto porque este índice é calculado em função da área e do perímetro, porém se pode observar áreas onde ainda havia uma certa confusão, e para uma maior discriminação destas áreas ainda confusas, com o auxílio da razão entre os eixos e o 1º momento que, como a compacidade, são invariantes à rotação, translação e mudanças de escala, os resultados obtiveram um aumento na acurácia.

Os outros índices de forma utilizados não foram tão relevantes, já que os mesmos nem sempre apresentavam a invariância quanto à rotação, translação e mudança de escala dos objetos. Isto dificulta a separabilidade entre segmentos de mesma classe por exemplo, mas com tamanhos ou em posições distintas.

Para um aumento da acurácia da classificação, sugere-se além de uma melhoria dos algoritmos de segmentação, a possibilidade da implementação de um *software* com a possibilidade da classificação *orientada ao objeto* utilizando-se do sistema das redes neurais artificiais, com a opção de manipulação de todos os parâmetros da arquitetura da rede, como é permitido do ambiente *MATLAB*. Porém neste, o tempo de processamento é considerado razoável, em comparação ao *software eCognition*, e com a implantação em uma linguagem mais apropriada este problema seria solucionado.

Considerando os ganhos obtidos com a integração da forma no processo de classificação, recomenda-se a pesquisa do uso de redes neurais para classificar

segmentos utilizando outras variáveis espaciais, como a textura e o contexto entre segmentos.

REFERÊNCIAS BIBLIOGRÁFICAS

ALBUQUERQUE, M. P.; ALBUQUERQUE, M. P. Processamento de Imagens: Métodos e Análises. **Revista de Ciência e Tecnologia**. (ISSN 1519-8022), v.1, n. 1, p. 10-22, FaCET 2000.

AVERY, T. E.; BERLIN, G. L. **Fundamentals of Remote Sensing and Airphoto Interpretation**. Prentice-Hall. New Jersey, U.S.A., 1992.

AZEVEDO, F.M.; BRASIL, L.M.; OLIVEIRA, R.C.L. **Redes Neurais com Aplicações em Controle e em Sistemas Especialistas**. Visual Books. Florianópolis, SC., 2000, 401p.

BÄHR, H. P. **Procesamiento Digital de Imágenes: Aplicaciones en Fotogrametria y Teledetección**. Hans-Peter Bähr Editor, 1991.

BÄSSMAN, H.; KREYSS, J. **Bildverarbeitung Ad Oculis**. Heidelberg Springer, 1998.

CENTENO, J. S.; TEDESCO, A.; ANDRADE, F. A.; MAFRA, C.; AUGUSTO, E. G. G.; SILVA, F. R. R.; LUZ, N. B. Análise da Forma em Imagens Temáticas. In: II Colóquio Brasileiro de Ciências Geodésicas - CBCG, Curitiba, UFPR, 2001. **Anais**. Resumo impresso, p. 204.

CHUVIECO, E. **Fundamentos da Teledetección Espacial**. Ediciones Rialp, S. A. Madrid, 1990, 449p.

CROSTA, A. P. **Processamento Digital de Imagens de Sensoriamento Remoto**. Gráfica da UNICAMP. Campinas, S. P., 1992, 170p.

DAHMER, A. **Segmentação de Imagens Ecocardiográficas Utilizando Redes Neurais e Medidas de Textura**. Porto Alegre, 1998. Dissertação de Mestrado - CPGCC/UFRGS.

DE COLA, L. Fractal analysis of a classified landsat Scene. **Photogrammetric Engineering and Remote Sensing**. v.55, n. 5, p. 601-610, 1989.

DUDANI, S.A. Aircraft Identification by Moment Invariants. **IEEE Trans. On Computers**. v.C-26, n. 1, p.39-46, 1977.

DYMINSKI, A S. **Análise de Problemas Geotécnicos Através de Redes Neurais**. Rio de Janeiro, 2000. Tese (Doutorado em Geotecnia) - Curso de Pós-Graduação em Engenharia Civil, Universidade Federal do Rio de Janeiro.

ECOGNITION. **Object Oriented Image Analysis**. Release 1.0. DEFINES AG rindermarkt. Munich, Germany, 2001.

ELSTNER, Rudolf. **Del Levantamiento de Catastro**. Mision Tecnica Alemana/ Catastro en CentroAmérica Y Panamá. Edicion Provisional, Noviembre,1971. San José, Costa Rica.

FABRI, J.A . Apostila de Redes neurais Artificiais. Conceitos e Aplicações. **IMESA**. Assis, S. P. , 1999.

FACON, J. **Processamento e Análise de Imagens**. Universidad Nacional de Córdoba, Argentina, 1993, 196p.

FITZGERALD, R. W.; LEES, B. G. Temporal Context in Floristic Classification. **Computers & Geosciences**. v. 22, n. 9, p.81-994, 1996.

FONSECA, L.M.G.; LOPES, E.S.S.; YAMAGUCHI, F.Y.; VINHAS, L. Apostila de Processamento Digital de Imagens. **INPE**. 2000.

FRANKLIN, S.E.; WILSON, B.A A Three Stage Classifier for Remote Sensing of Mountain Environment. **Photogrammetric Engineering and Remote Sensing**. v. 58, n. 4, p. 449-454, 1992.

FU, K.S.; MUI, J.K. A Survey on Image Segmentation. **Pattern Recognition**. n. 13, p.3-16, 1980.

GALO, M. L. B. T. **Aplicação de Redes Neurais Artificiais e Sensoriamento Remoto na Caracterização Ambiental do Parque Estadual do Morro do Diabo.** São Carlos, 2000. Tese (Doutorado em Ciências da Engenharia Ambiental) – Escola de Engenharia de São Carlos, Universidade de São Paulo.

GARCIA, G. L. **Sensoriamento Remoto: Princípios e Interpretação de Imagens.** Editora Nobel, 1982.

GERMAN, G. W. H.; GAHEGAN, M. N. Neural Network Architectures for the Classification of Temporal Image Sequences. **Computers & Geosciences.** v. 22, n. 9, p.969-979, 1996.

GONG, P.; HOWARTH, P. Frequency-Based Contextual Classification and Gray-Level Vector Reduction for Land Use Identification. **Photogrammetric Engineering and Remote Sensing.** v. 58, n. 4, p. 423-437, 1992.

GONZALES, R.C.; WOODS, R.E. **Processamento de Imagens Digitais.** Edgard Blücher Ltda. São Paulo, 2000, 509p.

GUYON, I. Neural Networks and Application Tutorial. **Physics Reports.** v. 207, p.215-259, 1991.

HAYKIN, S. **Redes Neurais. Princípios e prática.** Editora Bookman. Porto Alegre, 2001, 900p.

HEPNER, G. F.; LOGAN, T.; RITTER, N.; BRYANT, N. Artificial Neural Network Classification using a Minimal Training Set: Comparison to Conventional Supervised Classification. **Photogrammetric Engineering and Remote Sensing.** v. 56, p.469-473, 1990.

HU, M. Visual Pattern Recognition by Moment Invariants. **IRE Trans. Inform. Theory.** v. IT-8, p.179-187, 1962.

HUTCHINSON, F.C. Techniques for Combining Landsat and Ancillary Data for Digital Classification Improvement. **Photogrammetric Engineering and Remote Sensing**. v. 48, n. 1, p. 123-130, 1982.

JARVIS, C.H.; STUART, N. The Sensitivity of a Neural Network for Classifying Remotely Sensed Imagery. **Computers & Geosciences**. v. 22, n. 9, p.959-967, 1996.

KHOTANZAD, A.; LU, J. Distortion Invariant Character Recognition by a Multi-Layer Perceptron and Backpropagation Learning. **IEEE International Conference on Neural Networks**. v. 1, p.625-632, 1988.

KUO, C.; CHUO, T.; LEE, R. Identification of Urban Characteristic using IKONOS High Resolution Satellite Image. In: 22 nd ASIAN CONFERENCE ON REMOTE SENSING, 2001. Singapore.

LEES, B. G. Neural Network Applications in the Geosciences: an application. **Computers & Geosciences**. v. 22, n. 9, p.955-957, 1996.

LIBERMAN, F. **Classificação de Imagens Digitais por Textura usando Redes Neurais**. Porto Alegre, 1997. Dissertação (Mestrado em Ciência da Computação) - Curso de Pós-Graduação em Ciência da Computação, Universidade Federal do Rio Grande do Sul.

LILLESAND, T.M.; KIEFER, R.W. **Remote Sensing and Image Interpretation**. John Wiley & Sons. New York, USA, 1994.

LIMA, R.F.P.; PHILIPS, J. A Importância do Cadastro Técnico Multifinalitário para o Desenvolvimento Econômico em Países Pobres. In: COBRAC 2000 - Congresso Brasileiro de Cadastro Técnico Multifinalitário, UFSC, 2000. **Anais**. Florianópolis, SC.

LOESCH, C.; SARI, S.T. **Redes Neurais Artificiais. Fundamentos e Modelos**. Editora da FURB. Blumenau, 1996, 166p.

MASCARENHAS, N.D.A; VELASCO, F.R.D. **Processamento Digital de Imagens**. Editora Kapelusz. Buenos Aires, 1989.

MARQUES, O F.; VIEIRA, H. N. **Processamento Digital de Imagens**. Editora BRASPORT. Rio de Janeiro, 1999, 406p.

MARTINS, M. P.; GUIMARÃES, L. N. F.; FONSECA, L. M. G. Classificador de Texturas por Redes Neurais.In: CBComp 2002 – II Congresso Brasileiro de Computação, 2002. **Anais**. Itajaí, SC.

MASON, D.D.; CORR, D.G.; HOGGS, D.C.; LAWRENCE, D.; PETROU, M. TAILOR, A M. The Use of Digital Map Data in the Segmentation and Classification of Remotely Sense Data. **International Journal of Geographical Information System**. v. 2, n. 3, p. 195-215, 1988.

MEYER, M. & WERTH. Satellite Data: Management Panacea or Potencial Problem? **Journal of Forestry**. v. 88, n. 09, p. 10-13, 1990.

MOREIRA, L.; ASSAD, E.D. Segmentação e Classificação Supervisionada no Uso da Terra na Microbacia do Córrego Lamarão. **Embrapa**. Brasília, 1998.

MURNION, S. D. Comparison of Back Propagation and Binary Diamond Neural Networks in the Classification of a Landsat TM Image. **Computers & Geosciences**. v. 22, n. 9, p.995-1001, 1996.

NELSON, M.M.; ILLINGWORTH, W.T. **A Pratical Guide to Neural Networks**. Addison-Wesley Publishing Company, Inc. USA, 1991.

NISHIDA, W. **Uma Rede Neural Artificial para Classificação de Imagens Multiespectrais de Sensoriamento Remoto**. Florianópolis, 1998. Dissertação (Mestrado em Engenharia de Produção) - Curso de Pós-Graduação em Engenharia de Produção, Universidade Federal de Santa Catarina.

PAOLA, J. D.; SCHOWENGERDT, R. A. The effect of Neural Network Structure on a Multispectral Land-Use / Land-Cover Classification. **Photogrammetric Engineering and Remote Sensing**. v. 83, n. 5, p.535-544, 1997.

PERELMUTER, G.; CARRERA, E.V.; VELLASCO, M.; PACHECO, M. A Reconhecimento de Imagens Bidimensionais utilizando Redes Neurais Artificiais. In: VIII SIBIGRAPI, 1995. **Anais**. Rio de Janeiro. p.197-203.

SANTOS, M.T.; ALMEIDA, A R.; TU, C.C.C. Redes Neurais Artificiais aplicadas no Monitoramento da Condição de Ferramentas de Corte utilizando Algoritmo de Extração das Características SFS. In: IV CONGRESSO BRASILEIRO DE REDES NEURAS, 1999. **Anais**. ITA, São José dos Campos, S.P. p. 292-297.

SCHNEIDER, W.; STEINWENDNER, J. Landcover Mapping by Interrelated Segmentation and Classification of Satellite Images. **Photogrammetric and Remote Sensing**. V. 32, part 7-4-3 w6, june, 1999.

SCHWARZ, M.; STEINMEIER, C.; WASER, L. Detection of Storm Losses in alpine Areas by Different Methodic Approaches using High-resolution Satellite Data. In: Bégni, G. Observing our environment from space: New solutions for a new millenium. Lisse, Balkema. p.251-257, 2002.

SONKA, M; HLAVAC, V.; BOYLE, R.. **Image Processing, analysis, and Machine Vision**. Brooks/Cole Publishing Company. USA, 1998.

TAFNER, M. A.; XEREZ, M. RODRIGUES, I.W. **Redes Neurais Artificiais: Introdução e Princípios de Neurocomputação**. Eko Editora. Blumenal, SC., 1996, 191p.

VARELLA, L. E. S.; PASSOS, E. P. L.; SANTOS, M. A.; ARAÚJO, R. L. Reconhecimento de Caracteres com Variações de Escala, Rotação e Translação. In: SIBIGRAPI, 2000. **Anais**. Rio de Janeiro. p.688-696.